The Dependency Distance Hypothesis for bilingual code-switching

Eva Duran Eppler
Roehampton University, London
e.eppler@roehampton.ac.uk

Abstract

This paper addresses the questions why and where, i.e. in which syntactic dependency relations, multilingual speakers are likely to code-switch. Code-switching (CS) is the linguistic behaviour of producing or comprehending language which is composed from lexical items and grammatical structures from two (or more) languages. This paper proposes that code-switching is more likely in syntactic relations with long dependency distances (Distance hypothesis DH). Dependency distance is the number of words intervening between a head and a depended. The DH is tested on a 93,235 word corpus of German/English monolingual and code-mixed discourse analyzed in Word Grammar (WG). This data set supports the DH in general and on specific syntactic functions. In ongoing work the DH is being tested on Welsh/English and Spanish/English corpora and with self-paced silent reading experiments using eye-tracking.

1   Introduction

This paper suggests that a property of dependency structures, i.e. dependency distance, accounts in part for syntactic code-switching. The idea that long dependency distances facilitate code-switching is an original proposal and can therefore only be indirectly linked to existing theories of code-switching.

The concept of dependency distance was first used in Heringer et al. (1980: 187) who call it 'Abstand'; the term 'dependency distance' was introduced in Hudson (1995: 16) who defines it as 'the linear distance between words and their heads, measured in terms of intervening words'. For an illustration of individual and mean distances see Figure 1
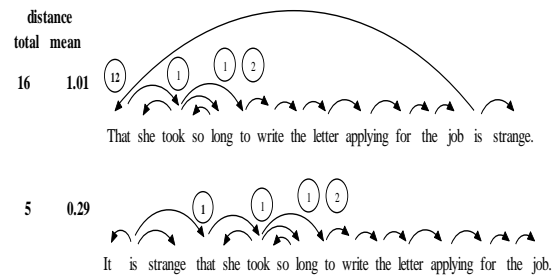

Figure 1 Dependency Distance

Mean dependency distances are cross-linguistically different (Liu 2008). In English most words are next to the word on which they depend (Collins 1996, Pake 1998). The main factor increasing distance is a change in dependency direction, i.e. a combination of left- and right-dependents (Termperley 2008, Hudson, personal communication).

Distance is an important property of a dependency relation because of its implications for the cost of processing the syntactic relation. Distance has been has been shown to be associated with syntactic memory load (keeping track of incomplete dependencies / obligatory syntactic requirements) and parsing complexity / integration cost (connecting a word into sentence structure) (Gibson 1998, Hiranuma 1999, Liu 2008). In terms of integration cost of long-distance dependencies, Gibson's (1998) Dependency Locality Theory (DLT) proposes that the longer a predicted category must be kept in memory before the prediction is satisfied, the greater the cost for maintaining that prediction. The greater the distance between an incoming word and the most local head or dependent to which it attaches, the greater the integration cost. In other words, the structural integration complexity depends on the distance between the two elements being integrated. That average dependency distance of a sentence can be used as a measure for its parsing complexity has been shown for centre-embedded vs. right-dependent sentences, subject vs. object relative

clauses[1] and Garden Path[2] sentences (Liu 2008).

Hiranuma (1999) demonstrated for Japanese and English that memory load increases with the length of a dependency, measured in terms of the number of intervening words. Liu et al. (2009) show that Chinese has a considerably longer mean dependency distances than other languages and propose that this may make Chinese more difficult to process than other languages. The average dependency distance of a text is an important comparative measure and can shed light on the cognitive demands of the language concerned relative to other languages (Liu 2008).

The present paper investigates the effects of dependency distance on syntactic code-switching, a linguistic phenomenon for which classical phrase-structure based models have proven to be unsatisfactory because over-generating (Eppler 0026).

2 The data

The study is based on a 93,235 words corpus of German-English monolingual and code-mixed discourse. The data is drawn from a community of Austrian Jewish refugees from the National Socialist regime who settled in London in the late 1930s. The L1 of the informants is Austrian German. The age of onset of the L2, British English, was during adolescence (15 - 21 years) for all speakers included in this study. At the time the audio-recordings were made (1993) all informants were in their late sixties or early seventies. A bilingual mode of interaction called 'Emigranto' developed among a close-knit network of community members. Linguistically the mixed code is characterized by frequent switching at speaker turn boundaries and heavy intra-sentential code-switching.

3 Dependency distance in English and German

English is generally considered to be a head-first language and allows for relatively little word order variation. As a consequence we get

few changes in dependency direction and short dependency distances. 63 − 74% (Collins1996 and Pake 1998 respectively) of English words that are syntactically related are also adjacent i.e. they have a distance of 0.

The mean distance between two syntactically related German words is hypothesised to be longer than the mean distance between two related English words. The main reasons why I assume German to have a longer mean distance are,
- the generally freer word order in German, which allows for more changes in dependency direction which trigger longer distances
- scrambling, i.e. word order variation of argument noun phrases with respect to each other (Example 1a & b), and/ or with respect to adverbial phrases (Examples 2) or even with respect to subjects (Example 3)

(1a)*Er hat ihr dieses Buch vielleicht gegeben.*
[3]%glo: he has her this book maybe given
(1b) *Er hat dieses Buch vielleicht ihr geben.*
%glo: he has this book maybe her given
(2) *Er hat ihr vielleicht dieses Buch gegeben*
%glo: he has her maybe this book given
(3) *dass jeder den Geruch erkennt*
%glo: that everybody this smell recognises

- the Verbalklammer, i.e. the discontinuity between AUX/MOD and main verbs
- different word orders in German main (V2) and subordinate clauses (V final or SOV).

4 Dependency distance in 'mixed' dependencies

'Mixed' dependencies are syntactic relations in which words A and B are from different languages. For mixed dependencies the main point of interest will be whether greater dependency distance influences / affects the chances of code-mixing. If code-switching is found to cause extra processing load, we might either expect
- shorter distances in mixed dependencies, because they 'counteract' the processing cost that is associated with code-switching (for some speakers)

---

[1] Processing cost of subject vs. object extraction, however, seem to be cross-linguistically different in English and German (Jackson and Dussias 2009)
[2] Late Closure (Frazier 1978) is preferred by the parser because it tends to minimize average dependency distance.

[3] CHILDES / LIDES transcription conventions are used throughout the paper
http://childes.psy.cmu.edu/manuals/chat.pdf
http://www.ling.lancs.ac.uk/staff/mark/lipps/easylides.htm

- a dependency distance between the mean distances for German and English monolingual dependencies, because syntactic dependency properties of both languages are involved
- longer mixed dependency distances, if we assume that the influence of a word's language on that of its dependent will decrease with increased distance. In other words, the longer the distance, the more likely we are to encounter an other language dependent, i.e. a code-switch. The latter assumption is similar to Gibson's computational motivation for the DLH. In an activation-based framework like WG, the head's activation will decay with distance (because of the limited quantity of activation in the system). The process of structural integration therefore involves reactivating the head to a target threshold level so that aspects of the head can be retrieved from memory. This reactivation is not only costly, but may also be incomplete.

For mixed dependency relations I am going to work on the hypothesis that the distance of mixed dependencies with a German head should be longer than the distance of mixed dependencies with an English head. This is based on the assumption that monolingual German dependencies are longer than English ones, and the possibility that heads influence dependency distance more than dependents. Furthermore, a change in dependency direction should be more frequent in mixed dependency relations with a German head, because verbs are main heads and they are involved in construction types like the Verbalklammer and V2 placement in main clauses and SOV placement in subordinate clauses.

The calculation of the mean distances in monolingual and mixed dependencies will reveal if these ideas are supported by the Emigranto data or not. The results on mean distances together with the standard deviation from the mean are presented in Table 1 Section 5.

5 General findings

Table 1 presents the mean dependency distances for monolingual German, monolingual English and mixed dependencies with German and English heads respectively.

	German	English	Average
Monolingual	0.87 (σ=0.78)	0.49 (σ=0.41)	0.68
Mixed with head	0.85 (σ=0.81)	1.26 (σ=1.08)	1.06

Table 1. Mean distances (and σ) in monolingual and mixed dependencies

These numbers tie in with pervious findings about dependency distance and the hypotheses formulated in Sections 1-4 as follows: Table 1 shows that:

1. monolingual German dependencies are longer than English ones. This supports the hypothesis made on the basis of the word order properties of the two languages (Section 3);

2. the mean distance of mixed dependencies with a German head is marginally shorter than the mean distance of monolingual German dependencies. This difference is unexpected but too small to support the idea that mixed dependencies counterbalance a potentially greater processing load for mixed utterances with a shorter dependency distance. This finding may, however, indicate that the word class that is assumed to increase dependency distance through a change in dependency direction, i.e. German verbal heads, is infrequently involved in mixed dependencies. Most importantly, however, it suggests that German words do not seem the affect the distance to their dependent, i.e. at least in terms of distance, they behave similarly in monolingual and mixed syntactic relations.

3. the mean distance of mixed dependencies with an English head is much longer than the mean distance of monolingual English dependencies. English heads thus seem to enter into 'looser', literally more remote, syntactic relations with German dependents. We would then expect English words to 'head' more dependency relations that are characterised by long distances, e.g. adjunct, extractee and extraposee relations. And we would expect German dependents of English heads to be more frequently located at the clause periphery. If we found more mixed dependents at the clause periphery in the Emigranto data, this would tie in nicely with the literature

on code-switching. Treffers-Daller (1994) first noted a high propensity of switching for 'dislocated constituents' in her French / Dutch data. Muysken (2000) subsequently adopted the idea that code-switching is favoured in peripheral positions as one of four primitives of code-switching.

4. the mean distance in mixed dependencies with a German head is approximately two thirds of the mean distance of mixed dependencies with an English head. This last finding completely contradicts the assumption that mixed dependencies with German heads are longer than mixed dependencies with English heads. This idea was based on the assumption that heads determine more characteristics of dependency relations than dependents, including the linear distance between them measured in terms of the number of words from one to the other.

5. the difference in mean distances between monolingual and mixed dependencies is highly significant ($X^2 = 18.6$, df = 1, p < 0.001);

6. The mean distance of mixed dependencies (1.06) is longer than that of both English and German monolingual dependencies. This finding supports the third possibility outlined above, i.e. that more distant words may have less influence on each other's language, because of the decay in activation as intervening words are processed and integrated into the structure of the input. If we assume that the influence of a word's language on that of its dependent decreases with increased distance, mixed dependencies may be the result of distance. By their very nature long distance dependencies in SVO and V2 languages are more likely to be located at the clause periphery. Treffers-Daller (1994) and Muysken (2000: 25) have both proposed peripherality as a factor favouring code-mixing.

7. and the standard deviation from the mean is higher for mixed dependencies. In other words, there is more variation in the distances of mixed dependencies and there are more mixed outliers.

These findings seem to suggest that the syntactic relations German heads enter with English dependents are not very different to the ones they enter with same language dependents, at least as far as distance is concerned. English heads, on the other hand, may enter into 'looser' and – literally – more remote (e.g. adjunct, extractee, extraposee) syntactic relations with German dependents. As a consequence, English words may function more frequently as heads of syntactic material that is located at the clause periphery.

The long dependency distances of mixed syntactic relations may furthermore point towards a processing motivation behind code-switching: the influence of a word's language on that of its dependent may decrease with increased distance. This would then mean that the longer the dependency distance, the more likely we are to encounter an other language dependent, i.e. a code-switch. This assumption, in combination with the findings presented in Table 1 discussed above, has led to the formulation of a claim about bilingual language use which combines features of grammar (syntactic relations) and psycholinguistics processes of speech production (dependency distance), the Distance Hypothesis.

Greater dependency distance of syntactic relations increases the chances of code-mixing.
(Eppler 2005)

The Distance Hypothesis is a syntactic processing hypothesis. Evidence in its support would therefore potentially shed light on both grammatical and psycholinguistics aspects of code-switching.

6   Specific findings

The analysis of individual syntactic functions in the Emigranto corpus statistically supports some of the constraints on code-switching proposed in the literature, but not others. The findings, for example, support the equivalence (Poplack 1980) and the subcategorization constraints (Bentahila and Davies 1983) in a probabilistic way. Both of these constraints are similar to the null hypothesis Eppler (2010) is based on, i.e. that each word in a syntactic dependency relation must satisfy the constraints imposed on it by its own language. The Complement Adjunct distinction (Mahootian and Santorini 1996: 470), on the other hand, is not supported.

The syntactic analysis of the Emigranto corpus moreover confirms that some syntactic functions are more easily switched than others.

Noun complements of determiners, for example, are clearly at the top of the 'borrowability' hierarchy; objects are more easily switched than subjects; and syntactically unrelated utterance elements are at the top of the 'switchability' hierarchy.

In the following three sub-sections, I will focus on the dependency distances of individual syntactic relations, comparing monolingual types with each other (Section 6.1), monolingual German ones with mixed ones with a German head (Section 6.2), and monolingual English ones with mixed ones with an English head (Section 6.3).

6.1 Monolingual German and monolingual English syntactic functions

Comparing monolingual German and English grammatical functions with each other on the one hand shows that German and English are typologically similar (they can be analysed with the same set of dependency types), but also reveals the main word order differences between the two languages. Sharers[4], objects, negatives, particles and prepositional are exclusively right-dependency relations of verbs in English. In German, sharers, objects, negatives particles and prepositionals of V2 verbs are also right-dependents, while they are left-dependents of clause final verbs. These results furthermore indicate that the German/English bilinguals possess two identifiable linguistic systems or languages, each with its identifiable grammatical rules and lexicon.

In Section 3 I outlined why I expect a longer mean distance for German dependency relations than for English ones, and we found this assumption confirmed by the data (Tables 1 and 2). The mean distance between two syntactically related German words is 0.87 in the Emigranto data; the mean distance between two syntactically related English words, on the other hand, is only 0.49. This is approximately 0.1 longer than what Hiranuma (1999) found for a comparable (conversational) corpus of 1,035 words, and closer to the 0.51 Liu et al. (2009) calculated for a written English sample text of about 100 words. Table 2 (Appendix B), however, shows that those monolingual English syntactic functions from the Emigranto

corpus that yield substantial enough a number of tokens to be included in the more fine grained analysis, have a mean distance of 0.4 and are therefore very close to Hiranuma's 0.386.

Tables 1 and 2 (Appendix B) therefore confirm that mean dependency distances differ cross-linguistically and that different dependency types have different mean distances (cf. Liu, Hudson and Feng 2009: 170). Table 2 furthermore reveals which grammatical functions differ most between German and English in terms of dependency distances. They are complements, subjects, sharers, objects and especially extractees.

The clause final placement of German finite verbs depending on complementizers could cause the longer mean distance of German complements: the head and the complement are literally at opposite ends of the subordinate clause. Subordination, however, tends not to be frequent enough in spoken language corpora to have this big an effect. The longer mean distance of German complements can be traced back to the same reason why we have significantly more German pre-adjuncts than English ones, i.e. attribute adjectives between determiners and nouns. The distance of German subjects (in their 'normal' position, i.e. as left-dependents) from their head verbs also deserves comment. The following two word order properties of German cause the, in comparison with English, longer mean distance. Firstly, the subjects of clause final finite verbs in subordinate clauses are almost at opposite ends of clauses. Secondly, material (e.g. adverbs) intervening between subjects and the verb that functions as the head / root of the sentences increases the distance of German subjects. Given that the Emigranto corpus contains a lot of Verbalklammern (because reference to past time is made with the present perfect rather than the simple past in spoken German), I find the mean distance for German sharers (1.64) relatively short, although it is of course three times longer than that of English sharers. The, for standard German, ungrammatically extraposed objects in the Emigranto corpus[5] shorten the mean distance of monolingual German sharers.

---

[4] 'Sharer' is a kind of verb complement. In other syntactic theories sharers are called xcomps or predicatives.

[5] Sub-stratum influence from Yiddish has rendered examples like these marginally acceptable in the Viennese dialect.

(4) hat mich gekostet zwei pfund zwanzig [*] .
%glo: has me cost   two pounds twenty

                    Jen1.cha, line 465

The mean distance for German sharers may also be a result of the monolingual German data containing more sharer relations between verbs and predicative adjectives than between auxiliaries / modals and non-finite verbs. Adjectives intervening between objects and their head verbs give rise to the longer mean distance of German object dependencies. The biggest difference in the mean distances between monolingual German and English dependencies, however, clearly lies in the extractees. An example that illustrates the 'damaging' effect of extraction (and the word order in subordinate clauses) on the mean distance of monolingual German extractees is (4)

*MEL: aber   wenn   man   einen   huegel hinauf#gehen muss -, das ist schon +...
%tra: and if one must walk up a hill, then that is already +...

                    Jen1.cha, line 447

Example (4) is a fragment, but the relevant syntactic relations are there and the extractee *wenn,* is six words away from its head, the main clause finite verb *ist*; the complement of the extractee (*muss*) is four words away from it; and the subordinate clause's subject (*man*) is three words away from its head. In extracted subordinate clauses that are not 'small' clauses, we get three changes in dependency direction between the words that build the basic syntactic structure of these clauses. This naturally increases distance.[6]

6.2 Monolingual German and mixed syntactic functions with a German head

Out of the most common syntactic relations (complements, subjects, adjuncts, sharers and objects), three show a significant difference between how often they occur monolingually and how often they enter mixed dependency relations with a German head in the Emigranto corpus. They are complements, subjects and adjuncts. The frequently switched

---

[6] Dependency distance can be quantified in different ways. Gibson, for example, quantifies it in terms of new intervening discourse referents. According to this measure we would expend 5 energy units when processing sentence (5).

complements are borrowed English nouns; subjects are infrequently switched, particularly subject pronouns like (5)

(5)
*LIL: **you** kannst # jauchzen .
%tra: you can   # rejoice

                    Jen2.cha, line 1019

despite linguistic constrains on switching subjects (Gumperz and Hernandez-Chavez 1971 and Pfaff 1975); and adjuncts a very frequently switched.

Table 1 showed that the mean distance of mixed dependency relations with a German head is actually a bit shorter than the mean distance of monolingual German dependencies (0.85 to 0.87). Table 3 (Appendix B), however, reveales that the distances for most mixed grammatical functions (subjects, adjuncts, pre-dependent sharers and post-dependent objects) are longer than their monolingual German equivalents. The slightly shorter mean distance of mixed dependencies with a German head (in comparison with monolingual German dependencies) is only attributable to three dependency types: complements, post-dependent sharers and left-dependent objects. Out of these three, it is the very large number of English complements with a German head, the borrowed English nouns, that brings the mean distance down.

This result also tells us something about the syntactic structure of mixed complement relations with an English noun: they are hardly ever pre-modified. A lot of the English predicative adjectives are also very close to their German head; and so are the English objects that depend on German clause final / SOV verbs. The fact that English post-dependent adjuncts are almost three times as far away from their German head as monolingual post-dependent adjuncts seems to support Treffers-Daller (1994), Mahootian and Santorini (1996) and Muysken (2000), i.e. that code-mixing is favoured in adjoined peripheral positions.

(6)
*MEL: nein # ich bin draussen # **as per usual**.
%tra: no # I am out

                    Jen2.cha: line 185.

In Section 5 we hypothesised that the mean distance of mixed dependencies with a German head might be marginally shorter than the mean distance of monolingual German dependencies because the word class that is assumed to increase dependency distance

through a change in dependency direction, i.e. German verbal heads, is infrequently involved in mixed dependencies. An analysis of all German verbs in the Emigranto corpus revealed that this word class does function as roots/heads in mixed dependencies.

A separate test performed on verb types (main vs. AUX/MOD) showed that overall German verbs are not significantly less frequently involved in mixed dependencies than monolingual ones (p=0.112). The same holds true for German main verbs (p=0.192). German auxiliaries and modals, however, are significantly more frequently involved in mixed dependencies than English ones (p=0.001). This finding is interesting as AUX / MOD are frequently in V2 position, which often coincides with the placement of verbs in SVO structures. German AUX and MOD are therefore placed in congruence sites (Sebba 1998). Congruence sites / equivalent surface word orders have been identified as factors that facilitate code-switching (cf. Muysken's four primitives of code-switching).

6.3 Monolingual English and mixed grammatical functions with an English head

In the Emigranto corpus five syntactic functions occur significantly more or less frequently switched with an English head than with both the head and the dependent from the English language. They are - again - subjects and (pre-)adjuncts, as well as sharers, extrapositions and extractions.

As for German, the corpus yields disproportionately fewer German subjects depending on an English verbal head than monolingual English ones, but they do exist. See (7)

(7)
*DOR: die **do-'nt mind** ## aber **I do** .
%tra: they don't mind ### but I do
                    jen1.cha: line 220.

The Emigranto data therefore provide probabilistic support for constraints on switching subjects or subject pronouns (Gumperz & Hernandez-Chavez 1971, Pfaff 1975).

Hardly any English verbs share their subjects with German words. This is unexpected and interesting for several reasons. For one, in this direction, i.e. h$_E$ d$_G$, we do not encounter the conflict in dependency

direction we get in this syntactic relation with a German head (where the dependents can be both, right-dependents or a left-dependents of clause final verbs). We would therefore expect switching to be 'easier' in this direction. Second, Hawkins (1986: 96) notes that German is much more resistant to sharer/xcomp structures than English and puts this down to the generally increased semantic diversity of basic grammatical relations in English. For the code-switched German / English data this means that it seems to be the semantics of the German dependent that constrains code-switching, not the English head.

The pre-adjunct relation, on the other hand, is very popular for switching between an English head and a German dependent among the Jewish refugees.

(8)
*LIL: die xx hat es # **in high heaven** gelobt.
%glo: xx has it #   praised
                    Jen2.cha, line 1570

(9)
*MEL: als kind **I didn't like anything**
          aber **I love food** .
%tra: as a child I didn't like anything
          but I love food
                    Jen2.cha, line 2058

Note that the pre-adjunct in (9) is also extracted, that is moved out of its default word order position and moved to the left clause periphery.

The difference between monolingual English and German extractees and extraposees depending on an English head is also highly significant. The next example illustrates a German long-distance (distance = 8) extraction.

(10)
*MEL: was die Dorit wieder geschmissen hat,
          I [/] **I would have liked**.
%glo: what the Dorit again   thrown   has,

It appears that for emphasis reasons speaker *MEL increases the distance of a mixed dependency relation from zero to eight in the above example.

The results presented in Table 4 (Appendix B), which compares the mean distances of monolingual English and mixed dependencies with an English head, strongly support the hypotheses formulated on the basis of Table 1 in Sections 1-5.

Hypothesis three proposes that English heads seem to enter into 'looser', literally more

remote, syntactic relations with German dependents. It furthermore predicts that we would expect English words to 'head' more dependency relations that are characterised by long distances, e.g. adjunct, extractee and extraposee relations. And we would expect German dependents of English heads to be more frequently located at the clause periphery (cf. Treffers-Daller 1994). This is exactly what we find in the data (see Table 4).

The Distance Hypothesis, states that greater distance seems to increase the chances of code-mixing. On the basis of Table 1 we assumed that the influence of a word's language on that of its dependent may decrease with increased distance, and mixed dependencies would therefore be the result of distance. As a consequence of their long dependency distances code-switches were also expected to be more frequently located at the clause periphery. This is again what we find in the data (see Table 4).

Focusing on the mean distances of individual syntactic functions in Table 4 we notice that ALL mixed dependencies with an English head (apart from objects) are longer than their monolingual English counterparts (this is unlike the mean distances of monolingual German and mixed grammatical relations with a German head (h_G) (Table 3)). Table 4 furthermore illustrates that all dependency relations that yield a significantly higher number of mixed tokens than monolingual ones (German adjuncts, extractees), are further away from their English heads than their English counterparts. The results presented in Table 4 therefore lend support to the finding that code-mixing is favoured in peripheral and adjoined positions.

The hypothesis that greater dependency distance of syntactic relations increases the chances of code-mixing therefore appears to apply particularly to mixed syntactic relations with an English head. Mixed grammatical functions with an English head seem to pose a particular processing complexity for the German/English bilinguals. The activation of English heads seems to decay especially rapidly in long-distance dependencies and render retrieving aspects of h_E, e.g. it's language, from memory particularly difficult. This appears to lead to the significantly larger number of mixed long distance syntactic relation with an English head in the Emigranto corpus.

7 Summary and Conclusion

The analysis of syntactic dependency relations in a 93,235 word corpus of German/English monolingual and code-mixed discourse analyzed in Word Grammar (WG) showed that the bilinguals possess two identifiable linguistic systems, each with its grammatical rules and lexicon. The code-switched speech results from the interaction between lexical elements and grammatical rules from these languages.

The data support some of the syntactic constraints on code-switching proposed in the literature in a probabilistic way: the constraint on switching subject (pronouns), the equivalence of structure (Poplack 1980) and the subcategorization constraints (Bentahila and Davies 1983). The Complement Adjunct distinction Complement Adjunct distinction (Mahootian and Santorini 1996: 470), on the other hand, is not supported, not even if we analyse the English noun complements of German determiners discussed in Section 6.2 as borrowings.

The most interesting finding to emerge from the comparison of monolingual and mixed syntactic relations (Table 1) in the corpus is that mixed syntactic relations have a longer mean dependency distance than monolingual ones. This led to the formulation of the Distance Hypothesis and a set of corpus-specific hypotheses on the syntactic behaviour of linguistic elements in Emigranto. The central syntactic processing claim to emerge from the quantitative analysis is that the influence of a word's language on that of its dependent appears to decay with the number of words intervening between it and its dependent. In other words, the longer the dependency distance, the more likely we are to encounter an other-language dependent, i.e. a code-switch.

The analysis of individual grammatical functions in Section 6 revealed that (with three exceptions) ALL mixed dependency relations are, on average, longer than the corresponding monolingual ones. In particular, the Emigranto corpus contains a considerable number of very long-distance mixed (post-)adjuncts with a German head, and English heads generally tend to enter into 'looser', literally more remote, syntactic relations with German dependents, i.e. syntactic relations that are not essential for building sentence structures, like adjunction, extraction (and extraposition). These

grammatical relations are also associated with long distances.

Including syntactically unrelated sentence element in the analysis, we can conclude that the ease with which elements are switched in the Emigranto corpus can be arranged on a continuum ranging from syntactic relations with very short dependency distances (such as subjects), to syntactically loosely connected grammatical functions with long dependency distances (such as adverbials, extractees and extraposees), to syntactically unrelated discourse elements at the utterance periphery (such as interjections, discourse markers and tags).

Dependency distance has been shown to play an important role in code-switching: syntactically related words are significantly more often in the same language when they are adjacent (Muysken's Adjacency Principle), and more distant words seem to have less influence on each other's language and are therefore more frequently switched. The Distance Hypothesis is an original proposal and can only be indirectly linked to existing theories of code-switching. It incorporates the idea that code-switching is favoured in peripheral (Treffers-Daller 1994, Muysken 2000) and adjoined positions but captures this notion on a more general syntactic processing level.

In collaboration with the Centre for Research on Bilingualism in Theory and Practice at the University of Wales, Bangor the Distance Hypothesis is being tested on other bilingual corpora (Spanish/ English and Welsh/English) and in self-paced silent reading studies supported with eye-tracking technology. In future work I also intend to investigate the effects of different kinds of 'interim' words, i.e. words intervening between the head and the dependent.

References

Bentahila, Abdelali, and Eirlys E. Davies. 1983. The Syntax of Arabic - French Code - Switching. *Lingua* 59: 301 - 30.

Collins, Michael John. 1996. A new statistical parser based on bigram lexical dependencies. *Proceedings of the 34th Annual Meeting on Association for Computational Linguistics. University of California, Santa Cruz 24-27 June 1996.* 184 - 191.

Gumperz, John J., and Eduardo Hernandez-Chavez. 1971. Cognitive aspects of bilingual communication. Language use and social change. In: W. H. Whiteley (ed.). Oxford: Oxford University Press 111 - 25.

Hawkins, John A. 1986. *A Comparative Typology of English and German. Unifying the Contrasts.* London and Sydney: Croom Helm.

Eppler, Eva. 2003. German/English database http://talkbank.org/data/LIDES/Eppler.zip

Eppler, Eva. 2006. Word Grammar and syntactic code-mixing research. In Sugayama, K. and Richard A. Hudson (eds.), *Studies in Word Grammar. New Perspectives on a Theory of Language Structure.* London, New York: Continuum, 117-39.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition,* 68, 1-76.

Heringer, H.J., Strecker, B. and R. Wimmer. 1980. *Syntax: Fragen-Lösungen-Alternativen.* München: Wilhelm Fink Verlag.

Hiranuma, So. 1999. Syntactic difficulty in English and Japanese: A textual study. *UCL. Working Papers in Linguistics 11* (1999): 309–322.

Hudson, Richard A. 2010. *An Introduction to Word Grammar.* Cambridge: Cambridge University Press.

Hudson, Richard A. 2007. *Language Networks. The New Word Grammar.* Oxford: Oxford University Press.

Liu, Haitao. 2008. Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science* 9(2): 161-174.

Liu, H., R. Hudson, Z. Feng. 2009. Using a Chinese treebank to measure dependency distance. *Corpus Linguistics and Linguistic Theory* 5(2): 161-174.

Mahootian, Sharzad, and Beatrice Santorini. 1996. Code-Switching and the Complement/Adjunct Distinction. *Linguistic Inquiry* 27: 464 – 79.

Muysken, Pieter. 2000. *Bilingual Speech. A Typology of Code–Mixing.* Cambridge: Cambridge University Press.

Pake, James. 1998. The Marker Hypothesis. A constructivist theory of language acquisition. PhD thesis, University of Edinburgh.

Pfaff, Carol. Constraints on language mixing: intrasentential code-switching and borrowing in Spanish/English. *Language* 55: 291 - 318.

Poplack, Shana. 1980. Sometime I'll start a sentence in Spanish y termino en Espanol: toward a typology of code-switching. *Linguistics* 18: 581 - 618.

Sebba, Mark. A congruence approach to the syntax of codeswitching. *International Journal of Bilingualism* 2(1): 1 - 19.

Temerpley, D. 2008.Dependency length minimization in natural and artificial languages. *Journal of Quarterly Linguistics* 15: 256-282.

Treffers-Daller, Jeanine. 1994. *Mixing Two Languages: French-Dutch Contact in a Comparative Perspective.* Berlin: de Gruyter.

Appendix A Notation – summary

Dependency types/ syntactic functions in WG	
Post-adjunct	*>a*
Pre-adjunct	*a<*
Complement	*c*

Pre-complement *(before 's, -ing)*	*c<*
Particle	*e*
Free complement	*f*
Indirect object	*i*
Negative *(not)*	n
Sharer/x-comp	*r*
Subject	*s*
Object	*o*
Prepositional complement	*p*
Extraposee	*>x*
Extractee	*<x*

(http://www.phon.ucl.ac.uk/home/dick/enc.html)

Appendix B – mean dependency distances for monolingual and mixed syntactic functions in Emigranto

	>c	s <	> s	> a	a <	> r	r <	> o	o <	> x	x <	> n	n <	> p	p <	total
G	0.65	0.54	0.07	1.1	0.37	1.64	0.07	0.78	0.83	-	2.16	0.33	0	-	0	0.73
E	0.22	0.07	-[7]	1.26	0.38	0.53	-	0.5	-	-	0	0	-	-	-	0.4

Table 2. Mean distances of monolingual German and English syntactic functions

	>c	s<	>s	>a	a<	>r	r<	>o	o<	>x	x<	total
G	0.65	0.54	0.07	1.1	0.37	1.64	0.07	0.78	0.83	-	2.16	0.73
hG	0.1	0.7	0.5	2.9	0.52	0.95	0.29	1.38	0.5	0.33	2.07	0.6

Table 3. Mean distances of selected monolingual German and mixed syntactic functions with a German head

	>c	s <	> a	a <	> r	> o	> x	x <	> n	Total
E	0.22	0.07	1.26	0.38	0.53	0.5	-	0	0	0.4
hE	0.84	0.9	1.33	0.78	2.12	0.18	0.45	3.5	-	1.05

Table 4. Mean distances of selected monolingual English and mixed syntactic functions with an English head

---

[7] For empty cells mean distances are not available.