Dependency distance and bilingual language use: evidence from German/English and Chinese/English data

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

Abstract

Closely related words tend to be close together in monolingual language use. This paper suggests that this is different in bilingual language use. The Distance Hypothesis (DH) proposes that long dependency distances between syntactically related units facilitate bilingual code-switching. We test the DH on a 9,023 word German/English and a 19,766 word Chinese/English corpus. Both corpora support the DH in that they present longer mixed dependencies than monolingual ones. Selected major dependency types (subject, object, adjunct) also have longer dependency distances when the head word and its dependent are from different languages. We discuss how processing motivations behind the DH make it a potentially viable motivator for bilingual language use.

1 Introduction

Corpus linguistic, computational linguistic and experimental language research has produced a considerable body of evidence over the past thirty years that there is a preference for linguistically related words to be close together in monolingual sentences (Gildea and Temperley 2010). Hudson (1995), Gibson (1998, 2000), Liu (2008) and others have worked on this from the comprehension side; Hawkins (1994, 2004), Temperley (2008) and collaborators have addressed the production side.

Most of this research captures the notion of linguistically 'closely related' and 'close together' with the concept of dependency distance/length. Dependencies are asymmetric syntactic relations between two words, a head/governor and a dependent. The head of each dependency is then the dependent of another word (unless it is the root of the sentence), forming a recursive structure which connects all the words of the sentence. Dependencies are (a) of a certain type, (b) directed, and (c) have a length.

(a) Dependencies can be semantic, morphological and/or syntactic. In this paper we are only looking at syntactic dependencies; the arrows representing dependencies are therefore labelled for grammatical functions, e.g. subject, adjunct etc.

(b) Dependency arrows point from the head to the dependent. Many languages have a dominant dependency direction: Arabic is predominantly head initial, Turkish head final; other languages like English, German and Chinese are more or less mixed.

(c) Every arrow spans a specific number of words (unless it indicates the root of the sentence). The linear distance between a head and its dependent, measured in terms of intervening words, is the dependency relation's distance[1] (Hudson 1995). The Mean Dependency Distance (MDD) of a sentence/text is the sum of its individual distances divided by the number of its dependencies.

Dependency distance (DD) is an important property of dependencies because of its implications for language production/processing. Constructing and interpreting sentences involves incrementally connecting words to arrive at meaning.

---

[1] Dependency distance can be quantified in different ways. Gibson (1998), for example, quantifies it in terms of new intervening discourse referents.

This process consumes human/ computational resources; it is 'costly'. DD has been shown to correlate with the cognitive cost of processing syntactic relations in terms of the memory resources required to keep track of incomplete dependencies (Gibson 1998, 2000; Hudson 2010: 279); and in terms of the cost of connecting a new/incoming word to syntactically related ones. The computational cost of integrating a word into sentence structure has been shown to depend on the distance between a word and the most local head or dependent to which it attaches (Dependency Locality Theory DLT, Gibson 2000). The DLT predicts that structures with longer dependencies are more difficult to process. It can account for a number of processing complexity phenomena, e.g. the relative ease of subject- vs. object-extracted relative clauses; ambiguity resolution in e.g. prepositional phrase attachment decisions, heaviness effects, and processing overload effects of multiple center-embedded structures.

Considerations of parsing complexity have also been proposed to affect language production (Hawkins 1994, 2004, Temperley 2008). Synchronically and on the level of the individual speaker this seems to manifest itself mainly in phenomena of syntactic choice, e.g. default word order vs. extraction/extraposition (Temperley 2008); diachronically Liu (2009) and Gildea and Temperley (2010) suggest dependency length minimization may also play a role in the shaping of grammars, i.e. language evolution.

As DD has implications for the cost of language processing, factors influencing dependency length need to be considered. Liu (2008) suggests that projectivity, no crossing arches in the dependency graph of a sentence, influences DD. Liu compared the MDDs of natural languages with those of artificial random languages, projective and non-projective ones. He found that non-projective artificial grammars have the longest MDD, followed by projective artificial languages and natural ones. Liu interprets his results as showing "the usefulness of a no-crossing approach to dependency length reduction" (Liu 2008: 14) and the reduced DD of natural languages as a consequence of projectivi-

ty (see also Gildea and Temperley 2010: 307). Most well-formed strings in natural languages are projective (Marcus 2007: 159).

Another factor that has been proposed to influence DD is dependency direction. If each word in a sentence has exactly one dependent, uniformly head-first or head-last structures yield shorter dependency distances than ones with pre- and post-dependents (Frazier 1985, Hawkins 1994, Rijkhoff 1994). Predominantly head-first or head last-languages, such as Arabic and Japanese, should therefore have the shortest MDDs. Liu (2010) has shown that this is not the case. The reason is that words can and do have more than one dependent.

If a word has more than one dependent, and the grammar requires all dependents to point in the same direction, and there is syntactic choice in terms of constituent order (e.g. a verb that has two prepositional dependents), placing the shorter dependent (phrase) closer to the head results in shorter dependencies. Hawkins (1994, 2004) reports that the preference of placing the shorter dependent closer to the head is found in head-first and head-last languages.

If a head has several dependents, placing all of them on the same side of the head creates a kind of 'crowding' effect. German subordinate clauses, which are head final (rather than V2), illustrate that all dependents of the verbal head (*haben*) crowd to its left.

(1)

I forgot, dass wir eine neue partie angefangen haben.
that we all again a new game started have
Jen1.cha, line 2541

In this case, there is no word order choice; if there was, placing some dependents to the left and some of the right of the verb would result in a shorter MDD. If a head has several dependents, balancing them on either side of the head results in shorter dependency lengths (Temperley 2008).

Languages that have a prevailing dependency direction but allow some short dependent phrases to branch in the opposite direction have shorter MDDs than consistently same branching languages

(Dryer 1992, Liu 2010). English is generally regarded as a predominantly head-first language. In the Penn Treebank, however, only 48.8% of the English dependencies are head-first; German was found to be, on average, 54.5% head-first and Chinese 31.5% (Liu 2010: 1571). Of the three languages we are looking at in this study, English has the best balance between left and right dependents and should therefore have the shortest MDD (followed by German and then Chinese). Section 5 presents empirical support for this hypothesis.

So far we have established that DD is an appropriate and widely used measure for establishing the linear proximity of linguistically related words. DD can therefore be used to test whether closely related words tend to be close together in monolingual and bilingual language use. Other properties of dependencies, the type of relationship they encode and their direction, were discusses as factors influencing dependency distance. Most importantly, the effect DD seems to have on the computational resources required for language processing and production was reviewed. Keeping track of long incomplete dependencies is a burden on memory load, and the cost of linking a new word into sentence structure - by connecting it to a head and/or dependent - also seems to be influenced by DD (Gibson 1998, 2000).
In the next section we will look at DD in the three languages involved in our data.

2   DD in English, German and Chinese

MDDs differ cross-linguistically. Although there is considerable variation in the type of language data analyzed to date (spoken, written; formal, informal) and ways in which distance is measured and calculated[2], there is a surprising amount of agreement as to which languages have short, and which ones have long DDs.

Out of the three languages we are looking at, we anticipated English to have the shortest MDD, followed by German and

Chinese. This is exactly what Liu (2008: 10) found: English (2.54) has a shorter MDD than German (3.35) and Chinese (3.66). Features of the three grammars that may account for this difference length will be looked at next.

The fact that English has fairly fixed word order and a prevailing dependency direction (head-first), but allows some short dependent phrases to branch in the opposite direction, seems to account for the short MDD of English (1.39, 1.49, 1.67 in Hiranuma's (1999), Eppler's (2010) and Wang and Liu's (2013) spoken data; 2.30 and 2.54 in the written data analysed by Gildea and Temperley (2010: 301) and Liu (2008: 12). In English, most words that are syntactically related are also adjacent; between 63% according to Collins (1996), 76% according to Pake (1998) and 78% according to Eppler (2010), but only slightly over 50% according to Liu (2008).

The mean distance between two syntactically related German words is longer than the mean distance between two related English words: 1.87 according to Eppler (2010), 3.07 according to Gildea and Temperley (2010), and 3.35 according to Liu (2008). The main reasons why German has a longer mean distance are the generally freer word order; the discontinuity between auxiliaries and their verbal complements (the Verbalklammer); and the different word orders in main (V2) and subordinate clauses (SOV). According to Liu (2008: 17) German has more adjacent dependencies than both Chinese and English.

Chinese has the longest MDD, not only of the three languages we are looking at in this paper, but also of the 20 languages Liu (2008) compared: 2.85 in spoken news data (Wang & Liu's 2013: 63), and 3.66 in written news data (Penn Chinese Treebank; Liu 2008: 12). The facts that Chinese has fewer mixed (head-first/head-last) dependencies than German and English and that Chinese is an isolating language, which marks e.g. tense, number and aspect with free (rather than inflectional) morphemes, has a significant influence on dependency length and the number of dependencies in a text.

---

[2] Eppler (2010) and Hiranuma (1999) measure dependency distance in terms of the number of intervening words; Liu (2008, 2009, 2010) in terms of the difference between the words' position numbers. Liu (2009) found the resulting difference in MDD to be small (1.81 vs. 1.89).

This brief cross-linguistic discussion of dependency length in English, German and Chinese has shown that rigidity of word order, consistency of dependency direction, and language type (isolating, inflecting) impact on a language's MDD. Collins (1996), Pake (1998), Eppler (2010) and Liu (2008) have looked into the relationship between dependency length and adjacency. These preliminary findings are difficult to interpret and more work needs to be done on this in the future.

The comparison of MDDs in different data sets furthermore supports the idea that DD is positively correlated with style (Liu et al. 2009: 171, Temperley 2008). Casual speech has shorter distances than more formal speech and writing, even when of the same genre (e.g. news, Liu 2009, Wang and Liu 2013). The average difference in dependency length between spoken and written data in English, German and Chinese is approximately one (1.02), with little variation between the three languages (Chinese 0.81, English 0.91 and German 1.34).

In the next sections we will look at bilingual data, speech which is constructed from lexical items and grammatical structures from typologically different languages (English and German, English and Chinese). We will test whether syntactically related words from different languages also prefer to be close together, or whether long dependency distances facilitate code-switching (DH); i.e. we will investigate the effects of DD on syntactic code-switching.

3 The data

The present paper is based on two bilingual corpora, a 9,023 word sample of a 93,235 word corpus of German/English (Eppler 2003), and a 19,766 word corpus of Chinese/English (Wang & Liu 2013). Both data sets were analyzed in the same dependency theoretic framework (Hudson 2007, 2010).

The German/English data was recorded in January 1993 among a close-knit network of members of the German-speaking Jewish refugee community who settled in London in the late 1930s. All speakers included in this sample are female and in their late sixties or early seventies. Their L1 is Austrian German. The age of onset of their L2, British English, was during adolescence (15-21 years of age) for all speakers. In informal settings like the ones recorded, the participants use a bilingual mode of interaction sometimes called 'Emigranto' (Eppler 2010). Linguistically this mixed code is characterized by frequent switching at speaker turn boundaries and heavy intra-sentential code-switching. The audio data were transcribed in the CHAT/LIDES (LIPPS Group 2000) format. The transcripts were manually annotated for word class, dependency type, dependency direction and dependency distance. See Table 1 for a summary of the data.

The Chinese/English data (Wang and Liu 2013) were audio-recorded from mainland China and Hong Kong TV or broadcasting programs from June to September 2011. Approximately 20% of the data are from interview programs; about 80% of the materials are news, social news, and entertainment news. Intra-sententially code-switched sentences were selected from the data, transcribed and syntactically annotated to build a Treebank containing the following information: linear position of the head and the dependent in the sentence, word class, language and a selected number of dependency types. The MDD of the corpus and of individual dependency types were calculated from the Treebank using formulae proposed by Liu (2009). See Table 1 for a summary of the data.

	German	English	Total	Chinese	English	Total
Word Tokens	5591	3432	9023	16267	3499	19766
Percentage	61.9	38.1	100	82.3	17.7	100

Table 1. Distribution of languages in the German/English and Chinese/English (Wang and Liu 2013) data

4 The Distance Hypothesis

In monolingual dependencies the head and the dependent are from the same language; in 'mixed' dependencies they are from different languages. The main point of interest for this paper is whether the MDDs of monolingual and mixed dependencies are similar or different. If they are significantly different, DD may have an effect on intra-sentential code-switching.

Theoretically the MDDs of mixed dependencies can be shorter, intermediate or longer than the MDDs of the monolingual

dependencies. They would, for example, be shorter, if code-switching consumed additional processing resources which are counterbalanced by the reduced processing cost of short dependencies (Gibson 2000). Dussias (2001) found that complexity and switch frequency are inversely related. As the German/English data is heavily switched, its production is expected to have incurred not additional resources. Mixed MDDs might be between monolingual means, because syntactic dependency properties of both languages are involved. Or they might be longer, if the influence of a word's language on that of its dependent reduces with increased distance. In activation-based frameworks the activation level of a word and its properties (e.g. its sense or language) will decay with distance. Comprehension studies have shown that structural integration involves reactivating a word to a target threshold level so that aspects relevant for its integration can be retrieved from memory. This reactivation is not only costly (Gibson 1998), but may also be incomplete; information about a word's properties may degrade partially or completely. If we assume similar processes to drive production[3], we may hypothesise that long dependency distances (DD $\geq$ 2) increase the likelihood of an 'other' language dependent, i.e. a code-switch (DH).

A decay in activation levels of syntactically related words over distance is assumed to be the motivating factor behind code-switching. Both the head and the dependent need to be - or be made - active at the point in language production when the dependency between them is being established. Activation levels of words (and their properties) decay as intervening words are being produced. In long dependencies the processing load is therefore high (Gibson 1998, 2000) and the priming effect between the head and the dependent low. Mixed dependencies/code-switches may result from long DDs because the influence the head and the dependent have on each other decreases with increased distance. The

DH is a syntactic processing hypothesis; evidence in its support would therefore shed light on grammatical and processing aspects of code-switching.

5 MDDs in the two corpora

The MDDs for monolingual and mixed dependencies in the German/English and Chinese/English copra are presented in Table 2.

	G	E	AVG	C	E	AVG
Mono	1.87	1.49	1.68	2.85	1.67	2.26
Mixed	1.85	2.26	2.06	3.54	2.81	3.18

Table 2: MDDs of monolingual and mixed dependencies with German, English and Chinese heads

The results for monolingual German and English support the hypothesis that monolingual German dependencies will be longer than monolingual English ones (made on the basis of the word order properties of the two languages in Section 2), and findings by Liu (2008) and Gildea and Temperley (2010).

The mean distances of mixed dependencies with a German head either indicate that heads do not have a more significant effect on dependency distance than dependents, or that German verbs, the word class that increases German MDD through bi-directional long-distance dependencies, are infrequently involved in mixed dependencies with a German head.

The mean distance of mixed dependencies with an English head suggests that English words enter into more remote syntactic relations with German dependents. We therefore expect a) English words to 'head' more dependency relations that are characterized by long distances (adjuncts, extractees and extraposees); and b) German dependents of English heads to be more frequently located at the clause periphery (Treffers-Daller 1994, Muysken 2000).

The highly significant difference between monolingual and mixed dependency distances ($p<0.001$) supports the idea that DD affects code-switching.

The recent analysis of a 19,766 word Chinese/English corpus (Wang and Liu 2013) supports the DH and has revealed interesting similarities and differences between the German/ English and Chinese/English data.

---

[3] Dussias' (2002: 98) study of the psycholinguistic complexity of code-switching revealed "a relatively clean convergence from [...] corpus analysis [which reflects production data] and on-line comprehension effects".

Chinese dependencies are longer than English ones (p<0.005). This was expected from the morphological and word-order properties of the two languages (Section 2) and supports findings by Liu (2008, 2009).

The MDD of mixed dependencies with a Chinese (L1) head and an English (L2) dependent is longer than that of monolingual Chinese dependencies (p<0.001; this is different to what we found in the German/English data), and also longer than the MDD of mixed dependencies with an English head and a Chinese dependent (p<0.05).

MDD increases more from monolingual English to mixed with an English head (+1.14) than from monolingual Chinese to mixed with a Chinese head (+ 0.69). This is similar to what we found in the German/English data, where the MDD between monolingual English and mixed dependencies with an English head increases by (+ 0.77); the MDD between monolingual German and mixed dependencies with a German head is virtually the same. Heads from the speakers' L1s (German and Chinese) therefore hold their dependents 'tighter' than L2 heads.

The mean distance of mixed dependencies with an English (L2) head and a Chinese (L1) dependent is also longer than that of monolingual English dependencies, but the difference is not quite as marked as in the German/English data (p<0.05 vs. p<0.001).

The average MDD of mixed dependencies (3.18) is longer than that of monolingual dependencies (2.26), and the average MDDs of mixed dependencies is longer than the MDDs of both English and Chinese monolingual dependencies (p<0.05).

In summary, the comparison of the MDDs from the German/English and Chinese/English data (Table 2) shows that
- MDDs are cross-linguistically different with English having the shortest MDD, followed by German and Chinese
- monolingual dependencies in mixed corpora are not significantly different to those found in comparable monolingual corpora
- the average MDDs of mixed dependencies are significantly longer than those of monolingual dependencies.

The patterns in the Chinese/English data (Wang and Liu 2013) largely correspond to those in the German/English data. Most importantly, greater DD also seems to increase chances of code-switching in Chinese/English bilingual speech. The findings from a typologically different language pair and data set therefore support the hypothesis that long DDs affect the language of dependents in that they render other language dependents more likely (DH).

6 MDDs of individual dependency types

In this section we will compare individual dependency types from the two data sets in terms of distance. The German/English data were analysed for the full range of syntactic relations used in Word Grammar (Hudson 2010). The Chinese/English data were analysed for four syntactic relations (subjects, objects, attributes and adverbials − both of the latter two are considered as adjuncts in the German/English data). To facilitate the comparison, we will focus on subjects, objects, and adjuncts. Section 6.1 briefly looks at monolingual dependencies, Section 6.2 compares monolingual L1 dependencies with mixed dependencies with an L1 head, and Section 6.3 compares monolingual L2 dependencies with mixed dependencies with an L2 head. The findings support the main idea outlined in the Section 5, the DH, and related findings from the code-switching literature (Treffers-Daller 1994, Mahootian and Santorini 1996, Muysken 2000).

6.1 Monolingual dependencies

Table 3 illustrates how individual dependency types contribute to the average DD of 1.87 for monolingual German and 1.49 for monolingual English dependencies in the German/English data.

	s <	> s	> o	o <	> a	a <	total
G-G	1.54 (142)	1.07 (45)	1.78 (54)	1.83 (36)	2.1 (100)	1.37 (86)	1.87 (754)
E-E	1.07 (130)	- (7)	1.5 (82)	- (0)	2.26 (72)	1.38 (44)	1.49 (596)

Table 3. MDDs and frequencies of selected monolingual German and English dependency types; s- subject; o- object; a - adjunct; left- (<) and right dependent (>)

The column entries of Table 3 demonstrate that different dependency types have different mean distances (Liu et al. 2009: 170); the rows show that MDDs differ cross-linguistically (Liu 2008, 2009)

and that the German/English bilinguals' grammars seem to be intact in terms of topological fields: there are no English left-dependent objects (nor x-comps). The MDDs that differ most significantly between German and English are subjects (and x-comps). These differences are caused by the subjects of clause-final finite verbs (which are at almost opposite ends of subordinate clauses) and the Verbalklammer. Gildea and Temperley (2010: 301) also found that verb position contributes to the longer dependency distances in German, but stress that it is not specifically the distance from subject to verb that results in this effect. Given that subjects tend to be short and can frequently be placed on either side of the verb in German (s< or >s), this finding is in line with the interrelation between dependency direction and distance discussed in Section 1.

The biggest difference in mean distances between monolingual Chinese and English also lies in the subject relation (p<0.001).

	S	O	Atr	Adv	Average
C-C	2.55 (940)	2.74 (849)	1.59 (1505)	2.45 (3039)	2.33
E-E	1.41 (130)	1.65 (91)	1.17 (296)	1.92 (104)	1.54

Table 4. MDDs and frequencies of four monolingual Chinese and English dependency types

Chinese prepositional constructions, such as *bei, ba, jiang* or *ge* and the complement of *di,* which are used as adverbials, must follow the subject but precede the modified verb; this increases the DD between the subject and the root of Chinese sentences, as in Example (2) where the DD between the subject *wo* and the verb *dang* is 3 in Chinese; the DD between *I* and *treat*, on the other hand, is only 1.

(2)   wo ba ta dang pengyou.
       I PREP him treat friend
        'I treat him as my friend.'

Wang and Liu (2013) found longer MDD of Chinese objects in comparison with English ones (p<0.001). Tense is realized by inflectional morphology in English; in Chinese tense is usually handled by function words which separate the object and the head. In Example (3) the DD between the object *book* and its head *bought* is 2. In Chinese, the DD between *mai* and *shu* is 4, because the complement of the classifiers *zhe* and *ben* and the perfect-

tense auxiliary *le* intervene between the object and its verbal head.

(3)   *wo mai le zhe-ben shu.*
       I buy AUX this-CL book
        'I bought the book.'

Examples like these raise the question what size unit DD should be measured in.

6.2 Monolingual L1 and mixed dependencies with an L1 head

Table 5 shows that in the German/English data the distances for most mixed syntactic relations (subjects, adjuncts and post-dependent objects) are longer than their monolingual German equivalents.

	s<	>s	>o	o<	>a	a<	total
G-G	1.54 (142)	1.07 (45)	1.78 (54)	1.83 (36)	2.1 (100)	1.37 (86)	1.87 (754)
G-E	1.7 (10)	1.5 (2)	2.38 (29)	1.5 (20)	3.9 (38)	1.52 (27)	1.85 (525)

Table 6. MDDs and frequencies of monolingual German and mixed dependencies with a G head

The slightly shorter mean distance of mixed dependencies with a German head is mainly due to the large number of borrowed English nouns complements of German determiners (> c: G-G MDD 1.65 (155) vs. G-E MDD 1.1 (309)).

That English post-dependent adjuncts are, on average, two words further away from their German head than monolingual German ones supports the notion that code-switching is favoured in adjoined peripheral positions (Treffers-Daller 1998, Mahootian and Santorini 1996, Muysken 2000), as in Example (3).

(3)
*MEL: ich bin draussen # **as per usual**.
%tra: I   am out
                        Jen2.cha: line 185.

The MDD of mixed adverbials with a Chinese head is also much longer than that of monolingual Chinese adverbials (p<0.001), see Table7.

	S	O	Atr	Adv	Average
C-C	2.55 (940)	2.74 (849)	1.59 (1505)	2.45 (3039)	2.33
C-E	2.7 (161)	2.85 (310)	1.48 (43)	5.65 (54)	3.17

Table 7. MDDs and frequencies of monolingual Chinese and mixed dependencies with a Chinese head

The Chinese/English corpus furthermore contains data that support the notion that code-

switching is favoured in clause peripheral positions. In Example (4) the English subject *fans* is dislocated from its default position (preceding *zhuanmen)* and moved to the left clause periphery. Because of extraposition and obligatory pre-posing of prepositional phrases before the verb, the distance between *fans* and its Chinese head *jisong* is 6; in the corresponding monolingual Chinese sentence, the distance between the Chinese subject *fensimen* and its head *jisong* is only 2.

(4) *Fans weile   xiang ta zhijing   zhuanmen*
       jisong xianhua.
   Fans in order to to him pay their respects specially
       posted flowers
   'In order to pay their respects to him, fans specially posted flowers.'
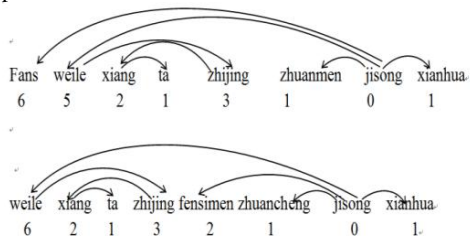
In Section 2 we suggested that the mean distance of mixed dependencies with a German head might be shorter than the mean distance of monolingual German dependencies because the word class that increases DD through a change in dependency direction, German verbal heads, is infrequently involved in mixed dependencies. An analysis of all German verbs in the German/English data revealed 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). German auxiliaries are frequently in the V2 topological field in German, a position that frequently coincides with SVO. German AUX/MOD are therefore placed in congruence sites (Sebba 1998). Congruence sites have been identified as facilitators of code-switching (Muysken 2000).

6.3 Monolingual L2 and mixed dependencies with an L2 head

In Section 5 we suggested that mixed dependencies may be the result of distance.

As a consequence of their long DDs, code-switches were expected to be more frequently located at the clause periphery in predominantly SVO and V2 languages.

More specifically, on the basis of the MMDs in the German/English data (Table 2) we proposed that English heads may enter into 'looser', literally more remote, syntactic relations with German dependents. We anticipated English words to 'head' more dependencies that are characterised by long distances, i.e. adjunct, extractee and extraposee relations, and predicted more German dependents of English heads to be located at the left or right clause periphery (Treffers-Daller 1994). This is what we find in the data.

	s <	> o	> a	a <	> x	x <	Total
E-E	1.07	1.5	2.26	1.38	1	2.3	1.64
	(137)	(82)	(116)	(116)	(1)	(3)	(596)
E-G	1.9	1.18	2.33	1.78	1.45	4.5	2.06
	(11)	(18)	(55)	(55)	(7)	(15)	(165)

Table 8. MDDs and frequencies of monolingual English and mixed dependencies with an E head; extraposee and extractee > x <

Table 8 demonstrates that all mixed dependencies with an English head (apart from objects) are longer than their monolingual English counterparts (this is unlike the MDDs of monolingual German and mixed dependencies with a German head; Table 6). Table 8 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. This finding supports the DH.

Table 8 shows that the adjunct relation is very popular for switching between an English head and a German dependent.
(5)
*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
The pre-adjunct in (5) is also moved out of its default word order position and extracted to the left clause periphery, which increases its DD by 4.
Example (6) illustrates a German long-distance clausal extraction.
(6)   *was die Dorit wieder geschmissen hat,*
          I [/] I would have liked.

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

It appears that for reasons of syntactic choice (Temperley 2008), speaker MEL increases the distance of a mixed dependency relation from zero to six in Example (6).

The hypothesis that L2 heads predominantly enter 'looser' longer syntactic relations with L1 dependents is also supported by the Chinese/English data, both numerically and in terms of DD.

	S	O	Atr	Adv	Average
E-E	1.41 (130)	1.65 (91)	1.17 (296)	1.92 (104)	1.54
E-C	2.75 (87)	2.88 (32)	1.67 (446)	2.07 (311)	2.55

Table 9. MDDs of four monolingual English and mixed dependencies with an English head

The MDDs of all mixed dependencies with an English head is longer than that of monolingual English dependencies of the same type, and there are significantly (p<0.001) more switched Chinese adjuncts (Atr and Adv) than subjects and objects, like in the German/English data. Note, however, that the increase in MDD between monolingual and mixed dependencies is bigger in subjects and objects than in adverbials. This may indicate that, if close syntactic relations are switched, their distance may have to be even longer. The distance between the Chinese subject *tamen* and its head is 2 in Example (7), in its English translation the DD between *they* and *send* is only 1.

(7)   *Tamen tiantian send E-mails.*
         They everyday send
          'They send E-mails everyday.'

In Example (8) the distance of *understand* and its object *yiqie* is 5; in its English translation the distance of *understand-everything* is 1.

(8) *I fully understand ni gaosu ta   de   yiqie*
                        you tell   him AUX everything
     'I fully understand everything that you tell him.'

The hypothesis that greater DD of syntactic relations increases the chances of code-switching appears to apply particularly to mixed dependencies with an L2 head. Mixed syntactic relations with an L2 head seem to pose a particular produc-

tion difficulty for the German/English and Chinese/English bilinguals alike, and the activation of L2 heads appears to decay more rapidly than that of L1 heads. This seems to render the retrieval of features of the L2 head (e.g. its language) from memory more difficult and lead to the significantly larger number of mixed long distance syntactic relation with an L2 head in both corpora. The findings from the German/English data presented in Table 8 and those from the Chinese/English Treebank (Table 9) also support the notion that code-mixing is favoured in peripheral and adjoined positions (Treffers-Daller 1994, Mahootian and Santorini 1996, Mysken 2000).

7   Summary and Conclusion

We started from a well-established principle of monolingual language comprehension and production: closely related words tend to be close together in the sentence. We then suggested that this may be different in bilingual language use, i.e. that DD has an effect on whether both words in a dependency come from the same language or not. The Distance Hypothesis proposes that long dependency distances increase the likelihood of an 'other' language dependent, a code-switch. This syntactic processing claim is based on the rationale that both the head and the dependent need to be active at the point in the production process when the syntactic relation between them is being established. If the head and the dependent are far apart, the influence (priming effect) of a word's language on that of its dependent will decay with time/distance (the number of words intervening between them). The longer the dependency link, the less the priming influence, and the more likely a change in language.

The analysis of a 9,023 word corpus of German/English monolingual and bilingual speech revealed that mixed dependencies have a longer MDD than monolingual ones. In a 19,766 word corpus of a typologically very different language pair, Chinese/English, mixed dependencies also have longer MDDs than monolingual ones (Wang and Liu 2013). The MDDs of both corpora thus support the DH.

The analysis of individual syntactic relations in both corpora revealed that, with

one exception in the Chinese/English corpus and three in the German/English data, all mixed dependency relations are, on average, longer than the corresponding monolingual ones. Both corpora contain considerable numbers of long-distance mixed adjuncts, and in the German/English data L2 heads predominantly enter into mixed long-distance syntactic relations that are not essential for building sentence structures (adjunction, extraction and extraposition). When L1 subjects and objects of L2 verbs are switched in the Chinese/English data, they have especially long dependency distances. In languages in which root verbs tend to occupy central sentence positions (SVO or V2), such as English, German and Chinese, long distance dependents will be located near the clause periphery. That code-switching is favoured in clause-peripheral positions has already been established in bilingualism research (Treffers-Daller 1994, Muysken 2000). The DH however, captures this notion on a more general syntactic processing level.

The results from the German/English and Chinese English data are promising. To establish DD between syntactically related units as a viable motivator for code-switching, the DH will have be tested on other bilingual corpora and with controlled psycholinguistic experiments to establish, e.g. that distance has similar effects in comprehension and production.

References

Michael J. Collins. 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.

Matthew Dryer. 1992. The Greenbergian word order correlations. *Language*, 68: 81-138.

Paola E. Dussias. 2001. Psycholinguistics complexity in codeswitching. *International Journal of Bilingualism* 9(1): 87-101.

John A. Hawkins. 1994. *A Perfomance Theory of Order and Constituency*. Cambridge University Press, Cambridge.

John A. Hawkins. 2004. *Efficiency and Complexity in Grammar.* Oxford Univerisity Press, Oxford.

Eva M. Duran Eppler. 2010. *Emigranto. The syntax of German/English code-switching*. Braumüller, Vienna.

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

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

Edward Gibson. 2000. The Dependency Locality Theory. In Y. Miyashita et al. (Eds.) *Image, Langauge, Brain*. MIT Press, Cambridge, MA, 92-126.

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

Richard A. Hudson. 1995. Measuring Syntactic Difficulty. Manuscript, University College London.

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

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

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

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

Peter Muysken. 2000. *Bilingual Speech.* Cambridge University Press, Cambridge.

James Pake. 1998. The Marker Hypothesis. Edinburgh University [PhD thesis].

Mark Sebba. 1998. A congruence approach to the syntax of codeswitching". *International Journal of Bilingualism* 2(1): 1-19.

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

Daniel Gildea and David Temperley. 2010. Do Grammars Minimize Dependency Length? *Cognitive Science* 34: 286-310.

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

Lin Wang and Haitao Liu. 2013. Syntactic Variation in Chinese-English Code-switching. *Lingua* 123: 58-173.