Adapting the traditional account of V2 using bilingual data

Emma Vanden Wyngaerd
Université Libre de Bruxelles

Abstract
The traditional Generative account of V2 does not al-low us to make any predictions for code-switched structures. In this paper, a grammaticality judgment task (GJT) was developed to test how native bilinguals judge code-switched sentences with a possibility for V2. The results of this GJT indicate that it is the finite verb which is responsible for word order. Some suggestions for how to incorporate this in the theoretical models are made. To be sure, extra investigation into word order in the subordinate clause is needed.

1. Background

Code-switching (CS) is one of many possible results of language contact. In fact, CS could be said to be "a hallmark of bilingual communities world-wide" (Poplack 2001: 2062). Studying the interaction of the two grammars involved in CS can be a valuable tool for investigating lin-guistic structure. It makes information accessible in a way which is not always available to us when looking at monolingual data only (for an example, see González-Vilbazo & López 2011).

Accounting for CS phenomena within the Generativist framework has been made easier since the advent of the Minimimalist Program (MP). In a framework such as X'-theory, in which the structure is built before lexical insertion, it is difficult to explain how exactly individual lexical items can influence the structure. The MP is a more "lexical entry driven" approach, and it is precisely the features on the items in the lexicon that drive the derivation (MacSwan 2009).

Adapting generative accounts developed for monolingual systems to explain bilingual data can be quite straightforward (see Cantone & MacSwan (2009) for DP word order). For other phenomena it is more complicated. A case in point is the V2 phenomenon, common in many Germanic languages, though not in English. In V2 languages, such as Dutch (1), the finite verb moves to the second position of the clause.[1] For Dutch, this movement can easily be discerned in sentences with a fronted adjunct, or sentences with compound tense, as Dutch is an SOV language (Koster 1975). In English (2), the finite verb remains in post-subject position. This is the case for most clauses as English does have V2 effects when the clause-initial constituent is either a *wh*-phrase or a negated phrase with scope over the whole clause. English is sometimes referred to as a 'residual V2' language (Rizzi 1990).

(1) Dutch: sov word-order, V2

a. Ik **zag** Adele.
b. Gisteren **zag** ik Adele.

[1] In subordinate clauses V2 word order does not surface in Dutch. This is usually attributed to the complementiser occupying C^0, blocking the movement of the verb. Some languages do have V2 in subordinate clauses, such as Yiddisch and Icelandic. For an overview of analyses of such cases, I refer to Holmberg (2015).

c. Ik **heb** Adele gezien.

(2) English: svo word-order, no V2

a. I **saw** Adele.
b. Yesterday I **saw** Adele.
c. I **have** seen Adele.

Traditional Generative accounts (such as the one in Holmberg (2015)) assume the following for V2 languages:

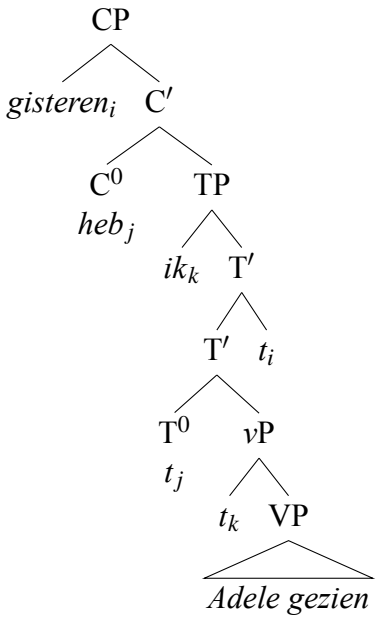
(3) a. a functional head in the left periphery (usually called C^0) attracts the finite verb[2]
b. this functional head then attracts something (which may be the subject, as demonstrated in (1a), or an adjunct as in (1b)) to the specifier position[3]

Non-V2 languages are thought *not* to have this functional head in the left periphery, at least not in declarative main clauses. This is shown in the simplified trees in (4). In Dutch, the finite verb has moved to C^0, prompting movement of the adverb *gisteren* to specCP. In English, on the other hand, no head is there to force movement, and the finite verb remains in T^0.

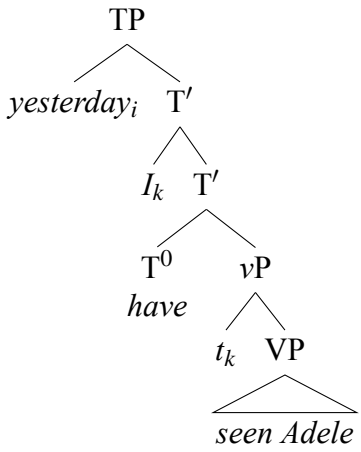
[2] This can also be one of the functional heads within a split CP analysis à la Rizzi (1997). See Westergaard & Vangsnes (2005) for an implementation.

[3] Note that Zwart (1993) handles this slightly differently; in his analysis, the verb *only* moves to C when a non-subject constituent is moved to specCP. Consequently (1a) and (2a) look the same underlyingly in this analysis.

(4) a.



b.



The prediction is then straightforward: if a CP layer is present, a V2 word-order will surface, if it is not, it won't. For a monolingual, the presence or absence of this projection will be learnt by exposure to the input. A Dutch-English bilingual, on the other hand, will acquire two systems: one *with* a C projection to trigger V2 word order (Dutch) and one without (English). So far, so good. However, if this bilingual then mixes the languages at his/her disposal, the traditional account makes no predictions. As it is a structural position – unlexicalised by any item – that determines the surface word order, this bilingual could choose either surface word order in (5).

(5) a. **Gisteren** *saw I Adele*.
b. **Gisteren** *I saw Adele*.

However, it has been clear from early studies that the grammar of CS is constrained Timm (1975). The purpose of the experiment described in this section is finding out what pattern native bilinguals prefer.

In the subordinate clause, where the C^0 is actually lexicalised, it does seem to be the case that it is the CP layer which determines the word-order. This has been investigated by Jansen *et al.* (2012), comparing

SOV with SVO languages. They conclude from their study of corpora of 18 German-Romance bilingual children that the C-head is indeed the determining factor for word-order effects in subordinate clauses:

> We have shown that the ordering of V and O can neither be systematically determined by the lexical non-finite verb in V nor by the finite modal/auxiliary verb in T, although both assumptions have been put forth in the literature. Furthermore, we have presented CS data indicating that the language of the complementiser seems to determine the underlying structural organisation within its phrasal complement, whereas the language of the finite verb in T is not relevant for its own position in TP. We have argued that all CS data presented here can be analysed in the same way, namely by assuming that the language of C determines the underlying syntax of the switched clause. (Jansen *et al.* 2012:370)

Other authors arguing for a similar influence of the C^0 over (subordinate) word order are Cantone (2007) and González-Vilbazo & López (2011).

As far as I know, there has only been one proposal in the literature that does not rely on an unlexicalised projection to trigger V2 word order. In Rambow & Santorini (1995), it is the finite verb itself which causes the movement out of the TP. It creates a SpecCP to which *something* must move. How or why exactly the verb moves in V2 languages, and not in non-V2 languages is left in the dark. Presumably, the causer of the movement is a feature on the verb. Hence, a Dutch verb would trigger movement to C, while an English verb would not. It is unclear, however, how this account would deal with residual V2 languages, such as English.

While not exactly relying on lexicalised projections, another interesting approach is the one developed in Westergaard & Vangsnes (2005). It was developed to deal with data from Norwegian dialects, which show very varied V2 patterns. They assume a split-CP (based on Rizzi 1997), in which each head in the CP layer may be specified separately for lexicalisation. If the specifier of a projection containing such a head is filled,

the head *must* be spelled-out by something, and it attracts the finite verb, resulting in a V2 order. As each head in the split CP is specified for "V2" separately, this account is well-equipped to deal with the dialectal variation of Norwegian.

It can also handle residual V2 languages deftly. For English, it would work like this. Th English CP, like all others, has many layers, amongst which there are the *wh*P and the negP. The heads of these projections are equipped with the feature that forces it to be lexicalised when the specifier position is filled. Consequently, when there is a negated phrase (6) or a *wh*-element (7) to fill the specifiers of these projections, V2 word order is triggered.

(6) a. As soon as **I arrived** at the station, the train came.
b. No sooner **had I** arrived at the station, than the train came.

(7) a. On Monday **that is** going to happen.
b. When **is that** going to happen?

While each of these are unlexicalised heads, their specifiers are filled, and one might then argue the language of the constituent in the specifier might determine whether or not the head will trigger V2. It is worth entertaining the possibility that it is indeed the language of the fronted constituent that determines the word order, when looking at the results from the GJT task.

2. Method

To investigate what word order native bilinguals prefer, a grammaticality judgment task (GJT) was devised.

2.1. Materials

All sentences were transitive and contained a fronted adjunct. The following factors were manipulated:

- word order: V2 or no V2
- language of the fronted constituent (FC)
- language of the inflected verb
- language of the direct object

Combination of these factors resulted in 12 conditions (2*2*2*2 = 16 - 4 monolingual conditions). Each condition was presented in three lexicalisations to improve statistical power, yielding the 36 test items. The survey also included 40 distractor items in CS mode.

Special care was taken to make sure that the switch point was unambiguous. Hence, no proper names occurred at the intended switch point. Switches between subject and verb were also avoided, as these are well known to be ungrammatical if the subject is a pronoun (first observed in Timm 1975).

In table 2.1, an instantiation of each condition is given.

		word order	language of the FC	finite verb	language of the object
1.	**In die winkel** *buys the dancer* **haar jurken**.	V2	Dutch	English	Dutch
2.	**Gisteren** *the boy ate* **een rijsttaart**.	no V2	Dutch	English	Dutch
3.	*Next year* **bezoeken wij het Rijksmuseum**.	V2	English	Dutch	Dutch
4.	*In the evening* **hij kuste zijn dochter**.	no V2	English	Dutch	Dutch
5.	*Sometimes forget students* **hun huiswerk**.	V2	English	English	Dutch
6.	*In the book Thelma draws* **een schets**.	no V2	English	English	Dutch
7.	**Wekelijks koken mijn ouders** *an Asian dish*.	V2	Dutch	Dutch	English
8.	**In maart Janne viert** *her birthday*.	no V2	Dutch	Dutch	English
9.	**Op de trein** *saw they a dog*.	V2	Dutch	English	English
10.	**Op zondag** *our son plays football*.	no V2	Dutch	English	English
11.	*In the garden* **vangt de kat** *a mouse*.	V2	English	Dutch	English
12.	*Every morning* **de priester zegent** *the holy water*.	no V2	English	Dutch	English

Table 1: Examples sentences per condition

2.2. Procedure

The grammaticality judgment task consisted out of an online questionnaire developed using the survey software Qualtrics. The survey consisted of the following parts, in the following order:

- Start-up screen. Contained a welcome message displayed in both Dutch and English.
- Language background questionnaire. Available in Dutch or English. Participants were able to choose which language to take the questionnaire in.
- Written instruction for the grammaticality judgment task, in code-switched text.
- Three trial items.
- Four randomised block of 19 items, randomly presented.
- Stimuli were pseudo-randomly distributed over the 4 blocks[4]
- Proficiency test. Ten items for Dutch and ten for English.
- Language attitude questionnaire. Available in Dutch or English. Participants were able to choose what language to take the questionnaire in.

Respondents were also asked to provide linguistic background information and answer questions about language attitudes. A small proficiency test was included. Results from respondents who scored badly on the proficiency test, and of later bilinguals[5] were excluded. If the language background questionnaire showed there was no longer daily use of both languages, responses were discarded as well.

In the instruction block, respondents were asked to rate code-switched sentences, on a seven point Likert-scale, ranging from 'completely unacceptable' to 'completely acceptable'.

[4] There was an equal representation of target and disctractor stimuli in each block. Conditions were more or less evenly distributed across the blocks.

[5] Acquisition of both languages before the age of four was used as a conservative cut-off point.

3. Results

3.1 Predictions

The factors were chosen to isolate the elements that are possible responsible for the word order.

- finite verb (as predicted by an approach along the lines of Rambow & Santorini 1995)
- language of the fronted constituent (as might be predicted by an adapted version of the approach by Westergaard & Vangsnes 2005)

In figure 1, the sentences are ordered according to expected grammaticality. As the language of the object is not predicted to influence word order, conditions 6 and 7 should be judged to be acceptable. These sentences have only a code-switched object DP, and word order matches the language of the rest of the sentence. Conditions 5 and 8 on the other hand present the mirror image and are expected to be judged unacceptable.

The interesting cases are between the two dotted lines. The ones above the double line represent conditions in which the language of the finite verb and word order match. The ones below represent conditions in which it the language of the fronted constituent and the word order match.

However, if the traditional generative account is correct after all, no such patterns should arise.

3.2 Results

A total of 100 responses were collected. Of those, 52 were discarded due to failure to complete the survey for unknown reasons. In total, 14 responses from early bilinguals were collected. Each condition had three lexicalisations, so a total of 42 judgments per conditions was collected.[6] Figure 2 shows the number of responses of each level on the Likert-scale per condition.

[6] Likert tasks are statistically well-powered at 30 judgments per condition (Sprouse & Almeida unpuplished ms:31).

	word order	language of fronted constituent	finite verb	language of the object
6.	no V2	English	English	Dutch
7.	V2	Dutch	Dutch	English
2.	no V2	Dutch	English	Dutch
3.	V2	English	Dutch	Dutch
10.	no V2	Dutch	English	English
11.	V2	English	Dutch	English
1.	V2	Dutch	English	Dutch
4.	no V2	English	Dutch	Dutch
9.	V2	Dutch	English	English
12.	no V2	English	Dutch	English
5.	V2	English	English	Dutch
8.	no V2	Dutch	Dutch	English

Figure 1: Overview of the conditions

The results are best analysed using non-parametric tests, as the different questions do not not come together to form a Likert-scale. Hence we are dealing with individual Likert items, which are ordinal data. They were analysed using χ -squared tests, as recommended by Boone & Boone (2012).

A one sample chi-square tests return a $p < 0.001$ for all of the categories (largest p-value = 0.0003415), meaning that the sentences are not rated randomly. This is no surprise, as we expect code-switched data to be as highly constrained as monolingual data.

In figure 2, we can see that the hypothesised preferences in 1 are impressionistically confirmed. The conditions above the double line (conditions 2, 3, 6, 7, 10, 11) seem to be outperforming the ones below it. Indeed, if we run a χ-squared test on the groupings of these categories, we find that they differ significantly ($p < 0.0001$). The effect holds up even if the the categories in the top and bottom two rows of the table are excluded ($p < 0.001$).

To see whether the language of the fronted constituent is a predictor of rating, sentences in which word order matched the language were grouped together. Testing whether there is a difference between a group-

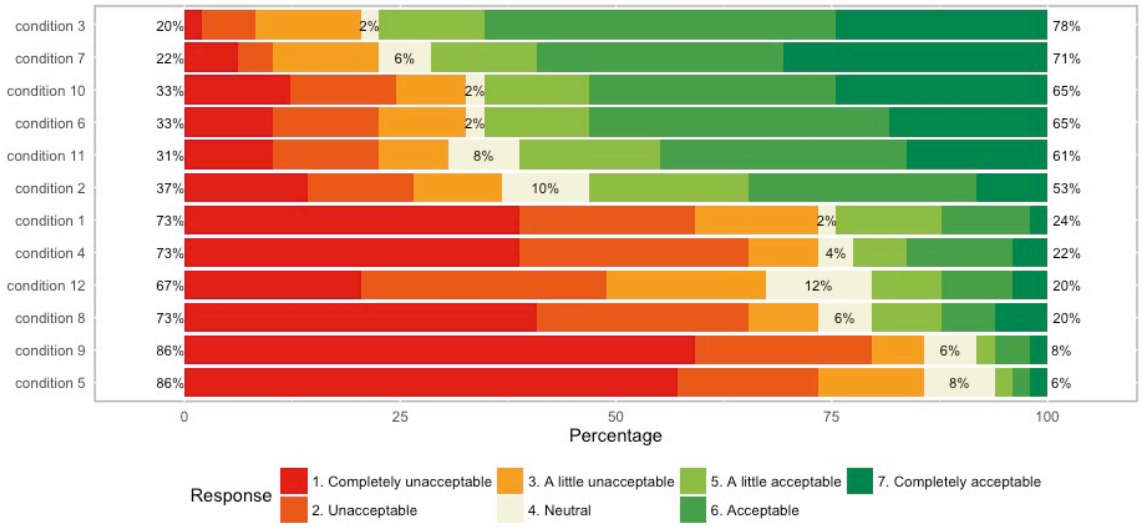


Figure 2: Results of the survey

ing of categories 1, 4, 6, 7, 9 and 12 versus the others does not yield a significant result ($p = 0.0847$).

There is no statistically significant difference between a grouping of categories 6 and 7 versus categories 2, 3, 10 and 11 ($p = 0.2437$). The same goes for groupings of categories 5 and 8 versus 1, 4, 9 and 12 ($p = 0.1444$). This suggests that the finite verb is solely responsible for grammaticality, rather than a combination of finite verb *and* fronted constituent.

A random grouping of categories (such as even versus odd or first six versus last six) also resulted in non-significant effects, with p-values of 0.1849 and 0.5476 respectively.

4. Conclusion and suggestions for further research

This grammaticality judgment task has confirmed that there is a strong preference for word order in code-switched sentences, contrary to what the traditional Generative accounts of V2 predict. The word order of the language that provides the finite verb is preferred. This is compatible with an analysis of the verb second phenomenon such as the one

sketched in Rambow & Santorini (1995). However, this analysis has not been very popular in the Generative literature, and it is quite divergent to many of the mainstream approaches. It remains a topic for further investigation how well this approach can be incorporated with the current state of the field.

These findings are also interesting, as they may be contrary to those in the few studies that have looked into word order differences in the subordinate clause. Jansen *et al.* (2012) and Cantone (2007) found the C^0 to be the projection responsible for word order. However, both of these studies have used the same corpora in a first language acquisition context. The diverging results may also be due to the different origins of the data; data from child language acquisition corpora are of a vastly different kind than data from grammaticality judgment tasks.

Data from this study should also be compared to investigations into word order in the subordinate clause and word order in main clauses with compound tense. This way, a comprehensive overview of the different factors determining word order should arise.

References

Boone, H. N. J. and Boone, D. A. (2012) 'Analyzing Likert Data'. *Journal of Extension* 50.

Cantone, K. F. (2007) *Code-switching in Bilingual Children*. Dordrecht, Springer.

Cantone, K. F. and MacSwan, J. (2009) 'Adjectives and word order: A focus on Italian-German'. In L. Isurin, D. Winford and K. de Bot, eds., *Multidisciplinary Approaches to Code Switching*. Amsterdam, John Benjamins Publishing Company, 243–277.

González-Vilbazo, K. and López, L. (2011) 'Some properties of light verbs in code-switching'. *Lingua* 121, 832–850.

Holmberg, A. (2015) 'Verb Second'. In T. Kiss and A. Alexiadou, eds., *Syntax – Theory and Analysis. An International Handbook*. 342-382, Berlin, De Gruyter. vol. 42 of *Handbooks of Linguistics and Communication Science*.

Jansen, V., Müller, J. and Müller, N. (2012) 'Code-switching between en OV and VO language: Evidence from German-Italian, German-French and German-Spanish children'. *Linguistic Approaches to Bilingualism* 2, 337–378.

Koster, J. (1975) 'Dutch as a SOV Language'. *Linguistic analysis* 1, 111–136.

MacSwan, J. (2009) 'Generative approaches to code-switching'. In B. E. Bullock and A. J. Toribio, eds., *The Cambridge Handbook of Lingusitic Code-switching*. Cambridge, Cambridge University Press, 209–335.

Poplack, S. (2001) 'Code-Switching (linguistic)'. In N. Smelser and P. Baltes, eds.,*International Encyclopedia of the Social and Behavioral Sciences*, 2062–2065.

Rambow, O. and Santorini, B. (1995) 'Incremental Phrase Structure Generation and a Universal Theory of V2'. In *Proceedings of NELS*. vol. 25.

Rizzi, L. (1990) 'Speculations on verb-second'. In J. Mascaró and M. Nespor, eds., *Grammar in progress: Essays in honour of Henk van Riemsdijk*. Groningen, Foris. 375–386.

Rizzi, L. (1997) 'The Fine Structure of the Left Periphery'. In L. Haegeman, ed., *Elements of Grammar*, Springer, 281–337.

Sprouse, J. and Almeida, D. (unpuplished ms) 'Power in acceptability judgment experiments and the reliability of data in syntax' Lingbuzz/001520.

Timm, L. (1975) 'Spanish-English Code Switching: El Porqué y How-Not-To'. *Romance Philology* 28, 473–482.

Westergaard, M. R. and Vangsnes, O. A. (2005) '*Wh*-questions, V2, and the Left Periphery of Three Norwegian Dialect types'. *Journal of Comparative Germanic Linguistics* 8, 117–158.

Zwart, J. W. (1993) *Dutch Syntax: a Minimalist approach*. PhD, Groningen, Rijksuniversiteit Groningen.