

RESEARCH

Focus particles and negative scope: Both evidence for syntactic integration?

Ira Eberhardt and Sam Featherston

University of Tübingen, DE

Corresponding author: Sam Featherston (Sam.Featherston@uni-tuebingen.de)

This article reports empirical work on the tests or judgement criteria which are used to determine the integration status of dependent clauses. Specifically it looks at focus particles and negative scope as indicators of such status in adverbial clauses. Some uses of these tests have produced conflicting results, which might suggest that they are not reliable criteria. Our empirical study suggest that both may be useful independently, but some factors may interfere with each other, producing falsified findings. Additionally, we observe that the test only works with certain sorts of focus particles, not with all, which could also potentially cause misleading findings.

Keywords: adverbial clauses; integration; experimental syntax; negatives; focus particles

1 Introduction

The relationships between clauses and between clause types in complex sentences are an issue which has attracted significant interest among syntacticians. In spite of considerable progress, there is still no complete understanding of the full set of relationships between the various classes of putatively dependent clauses and their matrix clauses. One essential tool which is required in order to carry out research in this field is some means of testing the degree of syntactic integration of one clause into another.

Some of the integration tests which have been advocated and applied are: correlatives, negative scope across a clausal boundary, focus particle scope, focus-background structure, the ability to occur in the pre-field (spec-CP) of a German sentence structure, (variable) binding across a clausal boundary, and others more. Since these tests are generally felt to concur, the consensus in the field seems to be that there is an over-arching factor of syntactic integration which the tests are accessing (e.g. Reis 1997; 2013, Haegeman 2003; 2012; Antomo & Steinbach 2010; Frey 2011).

But tantalizingly, the results of the tests do not always yield the same results, nor yield results of equal clarity, so that it is evident that we do not have a full understanding of their functioning (Pauly 2013; Christ 2014; Catasso 2018). In spite of this, there has so far been little exact investigation of their reliability and their range of applicability (but see von Wietersheim & Featherston 2019). This is thus a question on which it is apparent that more work could be done, in particular because it has on occasion been suggested that some of these tests deliver stronger evidence than others. A prime example of this is Pauly (2013: 274), who seems to consider negative scope the clearest indicator of integration, while some of the others afford less forcing evidence. There have also been differing views of the validity of some tests. For example, linguists have used placement

in the German pre-field as a criterion: if a sub-clause can occupy the spec-CP position of the matrix clause, then it is integrated. The pre-field has long been assumed to be a valid indication, but this has been recently questioned by Axel & Wöllstein (2009), Reis & Wöllstein (2010), and Frey (2011). Another criterion applicable to German is clause-final verb position. This has often been taken to be a reliable diagnostic (e.g. Speyer 2011), but this has also been questioned by Reis (2013). Coniglio (2011) assumes discourse particles to be a valid criterion, which is however questioned by Rapp (2018) and Grosz (to appear).

Our own intuitions too have provided us with examples where a test would be expected to yield a particular result, but it does not – on our judgements – do so. Sometimes these seem plausibly to be dependent upon fairly well-understood factors such as lexical choices, but other times no simple factor is apparent, and it is this which motivates the work presented in this paper. We hypothesize that some tests fail to work in certain contexts, and further, that there are cases where the nature of one type of test can block the successful functioning of another type of test.

In the light of this evidential uncertainty and our own interest in obtaining clear results on sentential integration as a part of our research effort, we decided to examine more closely, first, the combinability of certain tests, and second, some lexical choices apparently affecting tests. We therefore carried out two experimental judgement studies, gathering informants' intuitions of well-formedness in tightly controlled conditions. The results show that some tests are dependent on lexical and syntactic factors and show interaction effects when combined.

In the next section we will discuss in more detail the conditions for the tests and the issue of their combinability, which motivate our experiments. The third and fourth sections report the experiments that we carried out in order to address these questions. In section five we discuss the results and highlight their implications for linguistic work in this field.

2 Integration tests

A prime distinction which is made in discussions of the relationships between clauses is that of their degree of integration. As a first approximation, certain sorts of clauses, those which fulfil a syntactic or semantic role in the matrix clause, are also expected to be syntactically integrated into them, while other clauses, most usually those which do not stand in such a dependency relation to the matrix clause, are expected to be syntactically unintegrated.

We shall briefly present some of the criteria which are assumed to distinguish types of subordinate clauses and their different degrees of structural dependence here, basing our discussion on the proposals of Reis (1997) and Frey (2011). A major point in these papers is to establish that the distinction between integrated and non-integrated is not binary, rather we need to recognize at least a three-way distinction between integrated, relatively unintegrated, and absolutely unintegrated dependent clauses (Reis 1997: 128). We may use the integration criteria to determine their different structural behaviours.

The most essential distinction between those clauses labelled (absolutely or relatively) unintegrated and those that are integrated is that the latter occupy a structural position within the VP of their matrix clause. In (1) we understand that boredom is the reason why Vivian eats chocolate, and that both the behaviour and the cause are apparent. We can therefore analyse the causal modifier as a part of the matrix VP, attached just as a temporal or locational modifier would be. It is also within the scope of the evidential adverb *apparently*, which is outside the VP. This use is termed *propositional* (Sweetser 1990).

- (1) Vivian apparently [_{VP} eats chocolate [_{CP} because she is bored]].

There are however uses of causal clauses which seem to have a different relationship to the main clause – (2).

- (2) Vivian apparently [_{VP} eats chocolate] [_{CP} because her bin is full of chocolate wrappers].

Here the causal modifiers give the evidential basis for the whole proposition and are thus to be analyzed as relating to the evidential adverb. They are thus presumably attached above the VP. This use is termed *epistemic* (Sweetser 1990). The contrasting attachment points of these examples can be seen as the syntactic realization of their different semantic behaviour.

2.1 Variable binding

One criterion to distinguish integrated and non-integrated clauses is variable binding across the clause boundary (e.g. Reis 1997; Haegeman 2003; Reis & Wöllstein 2010; Coniglio 2011; Frey 2011; Christ 2014; von Wietersheim 2016). In an example with an integrated clause such as (3), the quantified phrase *every final-year student* can bind the variable in the dependent clause, so that we understand the pronoun *she* to pick up the preceding quantification.

- (3) Every final-year student_i has to study in the library [because she_i has exams next week].

This is not possible in unintegrated clauses. We have difficulty in assigning the variable interpretation to the pronoun so that it could be associated with the quantified phrase.

- (4) Every final-year student_i must be studying in the library [because she*_i isn't answering calls].

This difference is held to be the result of the different structural attachment points of the two clauses. The propositional *because* clause in (3) is thought to be attached to the verb phrase. In this low position, it is thus c-commanded by the quantified expression in the matrix subject position, permitting variable binding. The epistemic *because* clause in (4) is thought to be attached high as an adjunct to the CP of the matrix clause, so that no part of it is c-commanded by the matrix subject and variable binding fails. This would account for the difficulty of getting the bound reading in the second case. The success of variable binding can thus function as a marker of the integration of an adverbial clause.

2.2 Negation scope

Another example of a test of integration status involves the scope of negation (e.g. Haiman & Thompson 1984; Küper 1991; Wegener 1993; Günthner 1993; Haegeman 2003; Frey 2011). There are two different readings to an example such as that in (5).

- (5) Caroline does not like Draco [because he is a Slytherin].

On the first reading, Caroline does not like Draco and the reason is that Draco is a Slytherin. On the second reading, Caroline does like Draco, but not because he is a Slytherin. The second reading is available because the adverbial clause has a structural position within the matrix VP and it can therefore be in the scope of matrix negation.

Example (6) offers a contrast to the previous one.

- (6) Caroline does not like Draco [since he is a Slytherin].

Here only one reading is readily available, that where just the verb *like* is negated, so that we could rephrase it using *dislike*. The negation cannot easily scope over the *since* clause. The standard explanation of this would be that the *since* clause is not syntactically integrated into the matrix clause, a fact which would correlate with a higher attachment point, which would not be accessible to the matrix negation. The ability of a negative in the main clause to scope into an adverbial clause is thus regarded as evidence for the integrated status of an adverbial clause.

2.3 Focus particle scope

A third example test involves focus particles. These too have been proposed as a criterion for integration (e.g. Küper 1991; Haegeman 2012; Eberhardt 2017). We see this in examples (7) and (8).

- (7) Hannah likes to get up before sunrise [(especially/above all) if it's snowing].
- (8) Hannah likes to get up before sunrise [(**especially/*above all*) although it's snowing].

It seems quite acceptable to append a focus particle such as *especially* or *above all* to an *if* clause – (7), but not to an *although* clause – (8); the addition of a focus particle makes the sentence seem very unnatural. This is regarded as being a result of their different integration status, similar to the negation scope discussed above.

2.4 Problems with integration tests

These examples seem to show that the tests of integration status provide a consistent picture of what sorts of dependent clauses are integrated and what sorts are not. However, some other examples appear to reveal that the situation is more complex. In example (9-a) we see a case with negation scope into the clause, which is acceptable, as our description of the tests predicts, since this clause is an unambiguously integrated subordinate clause. However, the same example with a focus particle (9-b) seems unacceptable.

- (9) a. Alice is not coming [because you asked her to], but because there is work to be done.
 b. **Alice is not coming [especially because you asked her to], but because there is work to be done.*

This is surprising, because we had noted above that the addition of a focus particle to integrated clauses is unproblematic, in fact it is itself used as a test of integration. Example (9-b) is unacceptable, it is thus producing a false negative result: a linguist using these two examples to establish their integration status would thus come up with the wrong result. It would be tempting to suggest that the two factors negation scope in the subordinate clause and the placement of a focus particle on the subordinate clause are incompatible with each other. This account would capture our findings so far and it seems plausible: such incompatibilities do sometimes occur. However, the effect seems to involve yet more factors, since other examples do not show this interaction. For example, (10) seems very natural.

- (10) Alice is not coming [**only** because you asked her to], but because there is work to be done.

Deliberate or accidental use of two integration tests in one sentence may thus either lead to falsified results as in (9), or not as in (10). In any case, a cautious handling of the material can avoid the side effects of (9-b). Another example of a possible interaction of test material is from Coniglio (2011: 151), repeated here as in (11-a) and (11-b). Here the causal clause in (11-a) is classified as being integrated because the discourse particle *ja* (abbreviated as MP ‘modal particle’) is not possible when the adverbial clause is in the scope of the matrix negation. In contrast example (11-b) is classified as unintegrated because *ja* is acceptable.

- (11) Coniglio (2011: 151)
- a. Er ist nicht deshalb durchgekommen, weil er (*ja) schlechte
 he is not therefore come.through because he (MP) bad
 Noten bekommen hat, (sondern ...)
 grades received has (but.rather ...)
 ‘He did not succeed BECAUSE he (as you know) got bad grades, but because ...’
- b. Er ist nicht durchgekommen, weil er (ja) schlechte Noten
 he is not come.through because he (MP) bad grades
 bekommen hat.
 received has
 ‘He did not succeed because he (as you know) got bad grades’

The situation is however more complex: the unacceptability in (11-a) comes from the fact that discourse particles cannot be negated (Grosz accepted). The matrix negation scope is therefore not possible over adverbial clauses with a discourse particle. These considerations undermine the basis of the claim of non-embedded status for (11-b). First, the lack of matrix negation scope over the adverbial clause is possible in integrated as well as in non-integrated clauses. Second, the unacceptability of the discourse particle in the embedded clause in (11-a) does not indicate that its acceptability in (11-b) signals the root status of the clause. Indeed recent work has revealed that there is no simple mapping between discourse particles and the syntactic status of the clause (Rapp 2018; Grosz accepted). So the interaction of the discourse particle test with the negation test can lead to erroneous conclusions about the clause’s embedding status.

It was on the background of these doubts about the evidential status of these tests of integration that we decided to carry out an empirical investigation. There are in fact quite a number of other problems with integration tests that we could have addressed. For instance, the Spec-CP position test has been shown not to deliver the expected results in certain cases (Axel & Wöllstein 2009; Reis & Wöllstein 2010). The discourse particle test has also been claimed to be controversial (Grosz to appear). Positive results in tests using binding are not necessarily a result of grammatical binding but can be due to a weaker anaphoric relation (cf. von Wietersheim 2016). The experiments we report below are thus just a sample of the issues around integration tests which remain to be clarified.

3 Experiment 1

Our first experiment used German data, in part because these tests of integration are of particular interests to German linguists, as the integration status of certain dependent clause types is controversial and currently hotly debated (Reis & Wöllstein 2010; Frey 2011; Eberhardt 2017; Catasso 2018). We decided to utilize the negation and focus particle tests because they occur in contexts in which their scopes may (but also may not) overlap. In

other words, these tests provide an appropriate domain to examine a possible (negative) interaction of the integration tests using clauses which are uncontroversially integrated. Since the variable under debate is the perceived acceptability of the relevant examples, we chose to gather introspective judgements.

The experiment had two aims. First, we wished to see whether two integration tests would interfere with each other, selecting the integration tests scope of negation and focus particles. Our experiment materials thus contrasted structures in which a negative in the main clause scoped into a subordinate clause with other structures in which the negation was confined to the main clause. These structures were presented with and without focus particles on the subordinate clause. Thus far, therefore, the experiment had four conditions in a two by two design.

The second aim was to test whether various focus particles behave in the same way within this design. We thus tested four examples of focus particles: the exclusive particle *nur* 'only', the inclusive particle *auch* 'also', and the scalar inclusive particles *hauptsächlich* 'mainly', and *vor allem* 'above all'. These were intended first, to assure the generalizability of our results across lexical focus particles, but also to investigate whether the focus particles behave as a homogeneous group or not in these contexts, in the light of contrasts we noted above in (9-b) and (10).

This further test would extend our design to a two by two by four design. However, the factors used so far cannot be fully crossed, as the *no particle* conditions do not vary by particle. Technically, the choice of focus particle is nested within the conditions in which focus particles are present.¹ Our experiment thus contains the following variables and values:

- Negative scope into sub-clause: *not into sub-clause, into sub-clause*
- Focus particle on sub-clause: *no particle, particle*
- Focus particles: *nur, auch, hauptsächlich, vor allem*

The lexical material of the experiment consisted of a main clause and a propositional *weil* 'because' clause of a type that is generally taken to be syntactically integrated, so we would expect there to be no effects of the integration test criteria. If we observe differential results, this suggests that the factors tested are interacting with each other, or that the focus particles are not having the same effects.

We exemplify the experiment material in (12) to (15). Note that these four structural conditions are further differentiated by different focus particles. Here we use just *vor allem* 'above all' for illustration.

(12) *Condition 1: No scope into sub-clause, no focus particle*

Peter bleibt in seiner Wohnung. Er hat sie nicht gekündigt, weil die
 Peter is.staying in his flat he has it not given.notice because the
 Miete günstig ist, sondern er hat es nur angedroht.
 rent cheap is but he has it only threatened
 'Peter is staying in his flat. He has not given notice, because the rent is low, but
 he only threatened to.'

¹ Thanks to an anonymous reviewer for pointing this out. S/he suggested we use just two experimental variables, giving the variable *focus particle* the additional value of *no particle* to have ten conditions in a 2×5 structure. Although this is entirely correct, and we did repeat our experiments using this design, we did not finally adopt this approach in the data that we report here because it interferes with the statistics for our first aim, which requires equal numbers of data points in the four conditions. However we do use 5×2 design as a statistical test.

- (13) *Condition 2: No scope into sub-clause, with focus particle*
 Peter bleibt in seiner Wohnung. Er hat sie nicht gekündigt, **vor**
 Peter is.staying in his flat he has it not given.notice above
allem weil die Miete günstig ist, sondern er hat es nur angedroht.
 all because the rent cheap is but he has it only threatened
 ‘Peter is staying in his flat. He has not given notice, above all because the rent is
 low, but he only threatened to.’
- (14) *Condition 3: Negative scopes into sub-clause, no focus particle*
 Peter zieht aus seiner Wohnung aus. Er hat sie nicht gekündigt,
 Peter is.moving from his flat out he has it not given.notice
 weil die Miete gestiegen ist, sondern weil seine Mitbewohner rauchen.
 because the rent risen is but because his flatmates smoke
 ‘Peter is moving out of his flat. He did not give notice because the rent has risen,
 but because his flatmates smoke.’
- (15) *Condition 4: Negative scopes into sub-clause, with focus particle*
 Peter zieht aus seiner Wohnung aus. Er hat sie nicht gekündigt,
 Peter is.moving from his flat out he has it not given.notice
vor allem weil die Miete gestiegen ist, sondern weil seine
 above all because the rent risen is but because his
 Mitbewohner rauchen.
 flatmates smoke
 ‘Peter is moving out of his flat. He did not give notice above all because the rent
 has risen, but because his flatmates smoke.’

3.1 Expectations

The first two variables contain recognized tests of integration. Since the example sentences are types which are assumed to be integrated, the initial expectation would be that all four conditions produced by crossing these two variables will be rated as fully natural. Small differences between the conditions with and without negative scope into the sub-clause and the conditions with focus particles would not necessarily be problematic, since these are not identical structures. It would be unsurprising for an example sentence with a focus particle to be perceived to be more complex than one without, and this additional complexity could easily lead to lower ratings, since ratings are sensitive to processing and computational complexity.

If, on the other hand, the condition with the negation scoping into the sub-clause with a focus particle were rated less natural, we would have evidence of an interaction of the two integration test conditions, which could have the effect of the integration test being falsified. Our example (9-b) above seemed to hint at such a finding.

Our extended design also permits us to detect differences between focus particles. If the four focus particles did not produce parallel results, one might conclude that the integration test based upon them is more a lexical effect than one which is structural. The contrast between our (9-b) and (10) examples above opens the possibility of such a result revealing itself.

3.2 Procedure

Our chosen data collection method is called Thermometer Judgements (Featherston 2009). This method has been claimed to offer the most finely graded judgement data whilst simultaneously minimizing distortion (Featherston 2009). Informants carry out a sentence judgement task presented on a computer screen. They give introspective judgements of

example sentences after completing two training phases which accustom them to the task. Judgements are given on an anchored numerical scale. There are two anchor points, whose values are defined by the reference sentences in (16) (given for both English and German).

- (16) a. Reference example 1 is worth 20.
 English: The father buys the bread the children.
 German: Der Vater holt das frische Vollkornbrot ihm.
- b. Reference example 2 is worth 30.
 English: The father buys the children fresh bread.
 German: Der Vater holt dem Kind das frische Vollkornbrot.

While the first reference example is very unnatural and the second very natural, these do not anchor the end points of the scale available for the informants to use. Rather they act as non-terminal fixed points which informants can refer to while they give ratings on a linear scale. The aim is for most ratings to fall into the range of 15 to 35, in order to avoid the distortions which occur when numeric scores move from the two-digit range into the single-digit range, and even more so when they approach zero (Poulton 1989).

The informants' task is to judge the "naturalness" of the examples presented to them successively relative to the reference examples, which remain on the screen. The 16 experimental conditions and eight items were presented in four versions in a counterbalanced design, so that each person saw each item four times, and each condition twice. They were mixed with 15 filler sentences in a pseudo-random order produced by the experimental software package OnExp2 (Edgar Onea, Alexander Syring). A total of 32 native speaker informants participated in the experiment for a financial incentive, advertized at the University of Tübingen and made available online. The raw judgements were normalized to z-scores for visual presentation.

3.3 Results

The results are illustrated in Figures 1 and 2; the precise numbers are reported in the appendix. The charts show the ratings given on the vertical axis, transformed to z-scores. The error bars show the means and 95% confidence intervals around the means for each condition. The horizontal axis distinguishes the experimental conditions. Higher scores indicate that informants judged the structure to "sound more natural".

Figure 1 illustrates just the structural conditions, ignoring the different focus particles at this stage. The codings along the baseline indicate the structural conditions as follows:

- -NegScope: *no negative scope into sub-clause*
- +NegScope: *negative scope into sub-clause*
- -Part: *no particle*
- +Part: *with particle*

Figure 1 contains three error bars fairly low on the chart, and one somewhat higher. The higher bar is the condition with negative scope into the subordinate clause, but without a particle. This is something of a surprise, as the two structures without particles would be thought to be both fully acceptable. We had not predicted this, but it seems that the materials we created are perceived to be more natural when the negation has scope over the sub-clause instead of being restricted to the matrix clause. We find this plausible and invite the reader to re-read the material themselves. Even though the introductory statement (*Peter bleibt in seiner Wohnung/Peter zieht aus seiner Wohnung aus* ('Peter is staying in his flat'/'Peter is moving out of his flat')) tells the reader how to interpret the next part

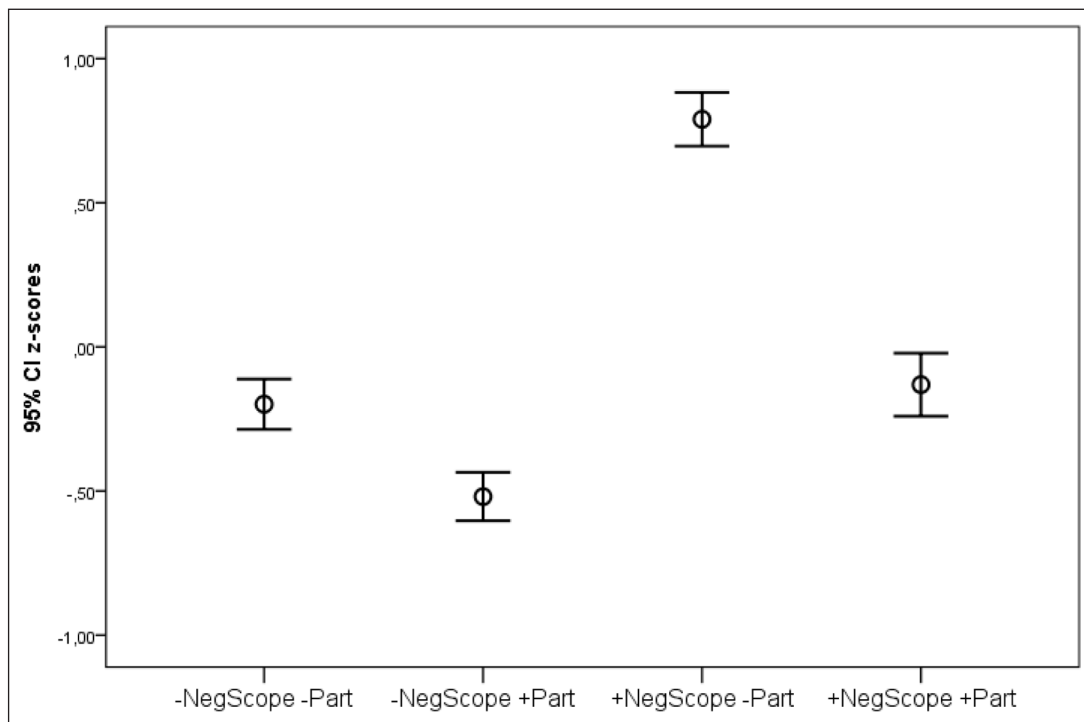


Figure 1: Experiment 1 results distinguishing the factors negative scope into subordinate clause and focus particle on the subordinate clause.

containing the negation (*Er hat nicht gekündigt ...*), the +NegScope conditions seem the more natural, perhaps because the continuation with *sondern* ‘but rather’ is better motivated in these cases. At any rate, while this finding was not predicted, it is not in any way a problem, because it is orthogonal to our tests and predictions. We just need to accept that in this sentence material, the cases where the negative scopes into the subordinate clause are perceived to be more natural.

The two conditions without particles have two functions. First, they allow us to see the basis of one of the tests of integration that we are focusing on. Since these subordinated clauses are integrated structures, the addition of negative scope into the subordinate clause is not expected to show any loss of acceptability, and it does not, quite the opposite, it actually improves the examples. But second, these conditions function as the baseline from which to observe the effect of the addition of focus particles onto the subordinate clauses. Here the picture is a different one: in the first two bars without NegScope there is a clear drop-off of acceptability with the addition of a focus particle, but this is no real surprise, since these sentences are a bit longer and a bit more complex. What is perhaps surprising is that this drop is much larger in the +NegScope case. The size of the loss in acceptability and the clear difference to the -NegScope case shows that there is a differential negative effect of applying a focus particle to the subordinate clause in this case. We tested this with a repeated measure analysis of variance using the factors +/-Negative Scope and +/-Particle, the interaction of which is significant: $F_1(1,31) = 30.60, p < 0.001$; $F_2(1,7) = 22.93, p = 0.002$. This simple 2×2 statistical test requires the same number of data points in all cells, which was the reason why we tested the full set of conditions without particles. The full data set is reported in the appendix.

So far we have not distinguished between the effects of the different focus particles, but we suspected that we might find such differences in the light of the apparent difference between (9-b) and (10) noted above. Figure 2 shows the same data set as Figure 1, but further discriminates by the four focus particles that we tested.

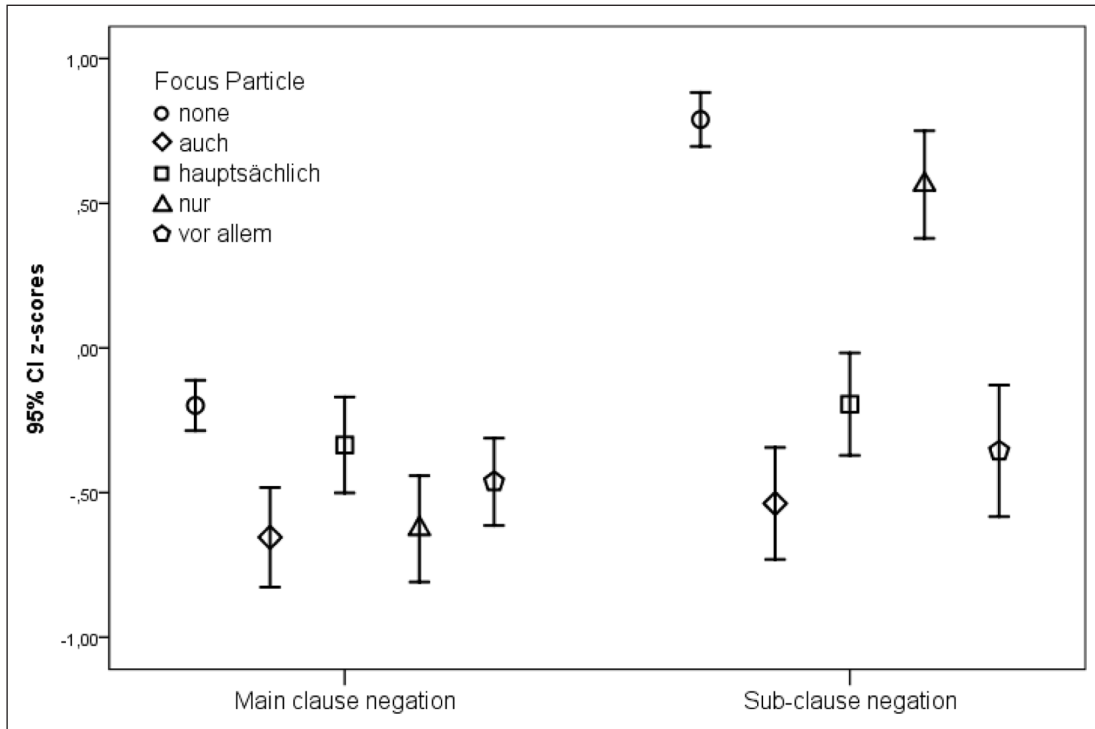


Figure 2: Experiment 1 results distinguishing the factors negation scope and focus particle presence as in Figure 1 above, but additionally distinguishing the different focus particles.

We shall look at the results from left to right. The five error bars on the left show the conditions without negative scope into the sub-clause, as before, but the +Part condition has been expanded to show the differences between the focus particles. We see that there are some fairly small differences between the bars representing the different focus particles: *hauptsächlich* and *vor allem* are a little better than the other two, but there is no reason why these focus particles should be judged identically.

Shifting our attention to the +NegScope conditions on the right-hand side of the graph, we see that the critical +NegScope +Part group is not homogeneous: while the focus particles *auch*, *hauptsächlich*, and *vor allem* are clearly worse than the +NegScope –Part error bar, *nur* behaves quite differently. It is much closer to the +NegScope –Part bar than to the other members of its own group. We tested this effect with a repeated measure analysis of variance, this time treating the lack of a particle as an addition value in the parameter Particle. It is therefore a 2×5 design for statistical purposes: the factor Negative Scope has the values +, – as before, the factor Particle has the values *auch*, *hauptsächlich*, *nur*, *vor allem*, no particle. In planned comparisons we tested whether the four focus particles are different to no particle in their interaction with Negative Scope. The focus particle *nur* is not different ($F_1(1,31) = 1.79, p = 0.19$; $F_2(1,7) = 1.79, p = 0.22$), the others are: *auch* ($F_1(1,31) = 29.43, p < 0.001$; $F_2(1,7) = 28.72, p = 0.001$), *hauptsächlich* ($F_1(1,31) = 30.49, p < 0.001$; $F_2(1,7) = 21.43, p = 0.002$), and *vor allem* ($F_1(1,31) = 30.18, p < 0.001$; $F_2(1,7) = 29.18, p = 0.001$). The apparent difference between *nur* and the others is thus confirmed.

3.4 Discussion 1

We have observed two phenomena of interest. First, we have found an interaction between the structural descriptions of some tests of integration: adding a particle to a –NegScope sentence has only a small effect, but adding one to a +NegScope sentence causes a drastic drop in acceptability. To this extent we have demonstrated that the integration tests can

interact. If a cautious linguist tries to apply two of them at the same time to be sure to obtain a clear result, the outcome may be the opposite of what they had hoped for. This result therefore warns us to be careful when applying these tests. Their effects are not necessarily consistent and independent of other factors; we should in particular be careful to choose very neutral example sentences to apply them to.

The second finding was that the focus particles did not all behave the same. While most of them showed the clear loss of acceptability in the +NegScope condition, *nur* did not. This would tend to confirm our intuitions on the examples (9-b) and (10) above, which differed only in the focus particle chosen. It is thus not sufficient to take care with negation in choosing what example sentence to apply the focus particle test to; a linguist also needs to choose the correct focus particle for the test to work. We discuss this in greater detail below.

There is one other aspect that we should note here and which relates to our perception of well-formedness. For the “bad” conditions – +NegScope +Part with *auch*, *hauptsächlich*, or *vor allem* – are in absolute terms no worse than the presumed “good” –NegScope conditions. Nevertheless, our judgements on the examples (9-a) and (9-b) above, which we repeat here as (17-a) and (17-b), were that there was a clear descent from (17-a) to (17-b) into something like unacceptability. We gave the supposedly unacceptable (17-b) a star, but our informants gave this type of sentence no worse ratings than the two –NegScope examples.

- (17) a. Alice kommt nicht, [weil du sie darum gebeten hast], sondern weil es zu tun gibt.
 ‘Alice is not coming [because you asked her to], but because there are things to be done.’
- b. *Alice kommt nicht, [**insbesondere** weil du sie darum gebeten hast], sondern weil es zu tun gibt.
 ‘Alice is not coming [**especially** because you asked her to], but because there is work to be done.’

This should make us think about the extent to which perceived acceptability is sensitive to the absolute acceptability of a sentence, and to what extent it is sensitive to the acceptability of one sentence relative to another. If we start from a +NegScope –Part condition, and add a particle, we will perceive a sharp reduction in acceptability. On the other hand, if we treated the –NegScope +Part condition as the starting point and extended the negative scope into the sub-clause, we should find no degradation, as the types –NegScope +Part and +NegScope +Part are rated the same. So whether we perceive that an example sentence is fully acceptable can depend upon what we choose to compare it with.

This has multiple implications for integration tests, which were thought to be fairly simple heuristics and are generally employed using the assumption of an absolute model of acceptability. In practice, when we gather introspective judgements with full control we see that acceptability is a continuum, not a binary distinction. We also generally see that the effects of priming or constraint violations are cumulative, not absolute. There is a lot to be said about this, but it cannot be addressed in detail here; instead we would refer the interested reader to Keller (2000) and Featherston (2005; 2009; 2019) where relative well-formedness in data and grammar are addressed in more detail. Instead we will move on to our next experiment which builds upon the first.

4 Experiment 2

After the very clear results of our experiment on German, we wished to extend our data base to another language in order to find out whether we would observe the same patterns. We therefore loosely translated our German material into English or constructed alternatives

where necessary – (18) to (21). The design of the experiment remained the same, as did the data collection method. The focus particles tested were *just*, *mainly*, *purely*, *specially*. We obtained our English speaking participants by having recruitment posters hung up in UK universities. In all 36 native speakers participated, but the data of four participants was discarded in order to obtain equal numbers in each of the four versions of the experiment. Two data points are missing, one because of a technical fault and one because it was implausible. The full set of materials and results can be inspected in the appendix.

- (18) *Condition 1: No scope into sub-clause, no focus particle*
Peter is staying in his flat. He did not terminate the lease, because the rent is quite low, but he threatened to.
- (19) *Condition 2: No scope into sub-clause, with focus particle*
Peter is staying in his flat. He did not terminate the lease, **just** because the rent is quite low, but he threatened to.
- (20) *Condition 3: Negative scope into sub-clause, no focus particle*
Peter is moving out of his flat. He did not terminate the lease because the rent had gone up, but because his flatmates smoke.
- (21) *Condition 4: Negative scope into sub-clause, with focus particle*
Peter is moving out of his flat. He did not terminate the lease **just** because the rent had gone up, but because his flatmates smoke.

4.1 Results

Our predictions and expectations were similar to those of experiment 1, only slightly modified by our first set of findings. We expect to see the same pattern of interference of the two test conditions. As before we first inspect the structural conditions, which are presented in Figure 3 ignoring the different focus particles. We may firstly note that the English results

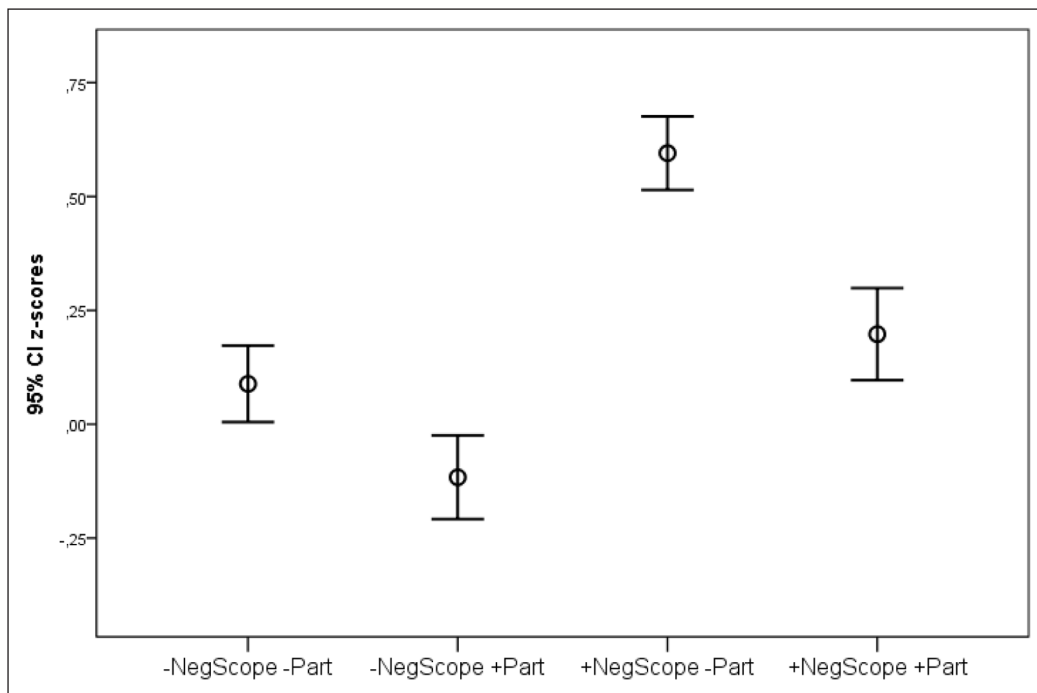


Figure 3: Experiment 2 results on English, distinguishing the factors negative scope into subordinate clause and focus particle on the subordinate clause.

show a similar picture to the German data. As before there is a small dispreference for adding a particle in the –NegScope condition and a larger one in the +NegScope condition. Nevertheless the interaction is significant in a repeated measure analysis of variance using the factors +/-Negative Scope and +/-Particle: $F_1(1,31) = 9.11, p < 0.005$; $F_2(1,7) = 10.09, p = 0.016$. This interaction may not at first sight appear a very strong effect, but we should recall that this interaction was actually only present for three of the four focus particles in the previous experiment – we shall see that it is here only borne by two – Figure 4.

The conditions are arranged here as in the previous experiment: on the left-hand side we see the –NegScope conditions. There is a bit more variation between the particles here in English: *mainly* is at least as good as the –Part condition, while *purely* is a bit weaker, and *just* and *specially* are clearly worse. This is no great surprise, as these conditions are not identical.

As before the +NegScope –Part condition is judged considerably better, and again there a much larger step down to some – but not all – of the focus particles in the +NegScope +Part condition. Cross-linguistic comparisons are difficult to make, but the impression is that the differences within the +NegScope conditions in English are smaller than in German.² Nevertheless, the drop is very clear.

We replicate here too the different behaviours of the focus particles in the +NegScope +Part condition. In German only one particle, *nur*, remained at the same height as the –Part condition. In English both *just* and *purely* do this, while *mainly* and *specially* become clearly worse. In planned comparisons we tested whether the four focus particles are different to no particle in their interaction with Negative Scope. The statistics confirm the

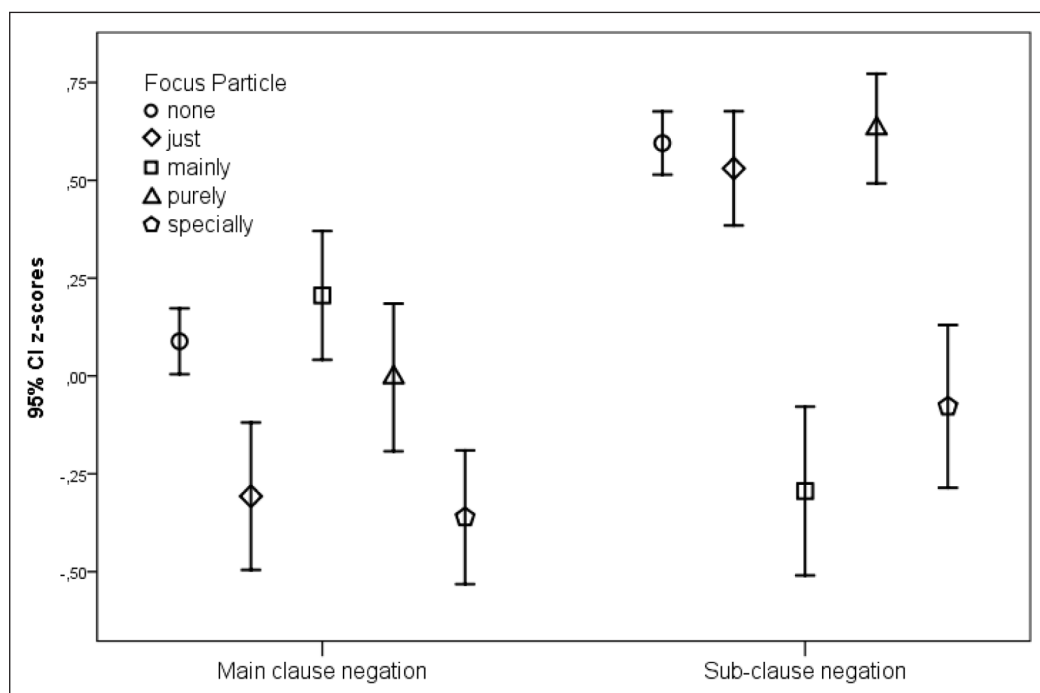


Figure 4: Experiment 2 results on English, showing the effects of negative scope and whether the subordinate clause is introduced by a focus particle, but additionally distinguishing the individual focus particles.

² Space does not permit us to discuss this in detail here, but both the experiments reported here included sets of standard items, which provide an absolute scale of perceived well-formedness (Featherston 2009; Gerbrich et al. 2019). Relative to these standard items the worst German sentences were quite unacceptable, while their English equivalents were still very much in the range of fully acceptable structures.

findings that we can observe, but it requires some thought to see why, so we shall relegate it to a footnote.³

4.2 Discussion 2

Our two experiments have shown very similar results; cross-linguistic confirmation of an effect is always a welcome sign that the generalization is robust. Our experimental studies have revealed a number of interesting findings. First, we have seen evidence that some tests of the integration status of subordinate clauses need to be treated with care. Above all the structural description of the negation scope test and the focus particle test seem to be incompatible with each other. All of the sentences tested here would be generally agreed to be syntactically integrated into their matrix clauses, and yet the application of the focus particle test to those structures with a matrix negative scoping into the subordinate clause would yield a result which would be thought of as identifying the subordinate clause as non-integrated. This is all the more problematic, because the base sentence which produced this result was apparently very natural – much more natural than its equivalent where the negation remains in the matrix clause. This finding thus joins the other work cited above which finds that the conventional tests of integrated status are less reliable than they are sometimes thought.

There is however a second issue which we have not yet discussed, namely, how we are to characterize the difference between the focus particles which do and those which do not react negatively to being applied to a subordinate clause into which there is negative scope from the matrix clause. The most obvious suggestion would be that exclusive focus particles (in our experiments: for German: *nur*; for English: *purely, just*) do not behave in the same way as the others in this regard. This would of course require further testing to be definitively established, but this would be a plausible generalization.

We are not aware of comments in the literature which have noted this effect, but it reminds us of the work on interpreting focus under negation of Neeleman & Vermeulen (2011). They divide the meaning contributions of focus particles into an A and a B component. So in (22), the A-component is *John invited Julia*, and the B-component is *John did not invite anyone else*. For *only*, the B-component is thus negative (which they term “polar”).

- (22) John invited only Julia
 a. [John invited Julia]
 b. $\neg\exists F', F' \neq \text{Julia}, [\text{John invited } F']$

Additive focus particles work differently. For example *even* in (23) has a positive B-component. So the A-component is *John invited Julia*, and the B-component is *John invited at least one other person who he was more likely to invite than Julia*.

- (23) John invited even Julia
 a. [John invited Julia]
 b. $\exists F', F' \neq \text{Julia}, [[[\text{John invited Julia}] <_{\text{likely}} [\text{John invited } F']]] \& [\text{John invited } F']]$

³ The focus particle *just* behaves differently to –Part ($F_1(1,31) = 7.34, p = 0.011; F_2(1,7) = 11.77, p = 0.011$), because it is judged worse than the –Part condition in –NegScope condition, but does not become any worse in the +NegScope case. On the other hand, *mainly* too shows a significant effect ($F_1(1,31) = 55.51, p < 0.001; F_2(1,7) = 36.76, p = 0.001$), but for the opposite reason: it was just as good as –Part in the –NegScope, but it drops off dramatically in the +NegScope case. The item *purely* does not differ systematically from no particle ($F_1(1,31) = 1.53, p < 0.224; F_2(1,7) = 1.08, p = 0.333$), it was about as good in the –NegScope case and remains so in the +NegScope case. Lastly, *specially* does not show an interaction either ($F_1(1,31) = 1.87, p < 0.181; F_2(1,7) = 4.57, p = 0.070$). It was already clearly worse than no particle in the –NegScope case, so the fact that it is much worse in the +NegScope case leaves it short of a significant interaction.

Their basic claim is that focus particles with “polar” B-components cannot appear in a negative context. Their formulation is in (24).

- (24) Material contained in the c-command domain of (local) sentential negation at LF cannot give rise to a polar B-component.

This predicts that we cannot say (25-a) but we can say (25-b), in a reading where the focus particle retains its narrow scope and does not move up over the negation.

- (25) a. *Sophie didn't invite only Martin.
b. Sophie didn't invite even Martin.

Intuitively, these facts discussed by Neeleman & Vermeulen (2011) are directly related to our phenomenon because of the necessary differentiation between exclusive and inclusive focus particles regarding their behaviour under negation. Even though our findings predict exactly the opposite – the exclusive focus particles behave better under negation than the inclusive ones – the incompatibility of our data and that of Neeleman & Vermeulen (2011) is perhaps only apparent. The mismatch may be due to the different negation types: sentence negation – as in (25), and constituent negation – as in (26). The constituent negation readings are not the only ones and perhaps not even the most accessible ones here.

- (26) a. Sophie invited not only Martin.
b. *Sophie invited not even Martin.

This is nothing like an explanation, but we suspect these phenomena and our experimental findings are linked: in each case the acceptability of negated examples depends on the semantic contribution of the focus particle (exclusive vs. inclusive). This may have deeper implications which we cannot follow up here and must leave to future research.

5 Conclusions

Our experiments have proved themselves well worthwhile because they have delivered a number of relevant findings. The background to our studies was the perception that tests of clausal integration are yielding inconsistent and sometimes even contradictory results. We noted in the text just a few of the places in the literature where this is apparent, for more detail see von Wietersheim (in preparation). This led us to seek the sources or causes of these inconsistencies, so this article can be seen as a methodological contribution to the debate on this linguistic issue. The paper is thus part of the wider trend in syntax research towards empirically validating the data basis of linguistic theory (e.g. Featherston 2007). This development is a response to the growing realization in the syntax community that even relatively small errors in the data basis can give rise to flawed theoretical conclusions.

We report here only a sample study, but even a small sample can motivate wider conclusions. The results first of all confirm that some well-formedness patterns assumed in the theoretical literature are valid and can be relied on. We would recommend further studies using such methods in this field where most papers are based only on the introspective judgements of individual linguists or corpus examples. While we do not reject these methods, we believe that more empirical work would be of value. One particular contribution would be that such studies could deliver clarification of some apparently conflicting claims in the literature. An additional benefit is to identify the precise patterns of acceptability brought about by integration tests and thus guide linguists in the application of such tests.

But second and more specifically, our experiments have demonstrated that the application of these tests is by no means as simple as might have been assumed. Certain tests are context-dependent and conflict with the structural descriptions of other tests. In our example, a linguist who used sentence negation in the matrix clause together with an inclusive focus particle test would receive falsified results. A similar unfortunate effect can be observed when using a discourse particle test in the context of negative (or other) scope. The generalization would be that example sentence material must be very carefully chosen because of the danger of side-effects: in this case, an exclusive focus particle can be employed, but an inclusive focus particle will lead to distortion or even a false result.

Our aim in this paper was to throw light on the empirical status of integration tests. In part, we have succeeded, but the additional data has opened up new questions that we previously did not even suspect existed. Like Goethe's Faust, we are forced to admit that our understanding is still very limited: but unlike Faust, we recognize that we have made progress, so we will not despair and turn to magic instead of scientific research.

Additional File

The additional file for this article can be found as follows:

- **Appendix.** Experimental materials. DOI: <https://doi.org/10.5334/gjgl.613.s1>

Acknowledgements

This work took place within project C6 *Expressions of Extrapositional Meaning: Synchrony and Diachrony*, project leaders Katrin Axel-Tober and Sam Featherston. The project belongs to the SFB 833 *Constructing Meaning*, funded by the Deutsche Forschungsgemeinschaft. Many thanks are due to the staff of the SFB 833 in general and our colleagues in project C6 in particular. Thanks too to Sarah Schwarz for her cooperation on the experimental work and to Hannah Gerbrich and Katja Bisinger for proofreading. Last of all thanks to an anonymous reviewer for very helpful comments.

Competing Interests

The authors have no competing interests to declare.

References

- Antomo, Mailin & Markus Steinbach. 2010. Desintegration und Interpretation: Weil-V2-Sätze an der Schnittstelle zwischen Syntax, Semantik und Pragmatik. *Zeitschrift für Sprachwissenschaft* 29, 1–37. DOI: <https://doi.org/10.1515/zfsw.2010.001>
- Axel, Katrin & Angelika Wöllstein. 2009. German verb-first conditionals as unintegrated clauses: A case study in converging synchronic and diachronic evidence. In Susanne Winkler & Sam Featherston (eds.), *The fruits of empirical linguistics 2: Product*, 1–35. Berlin: Mouton de Gruyter. DOI: <https://doi.org/10.1515/9783110216158.1>
- Catasso, Nicholas. 2018. Sind Sprechakte einbettbar? Evidenz aus dem Deutschen und Italienischen. In Elisabeth Leiss & Sonja Zeman (eds.), *Die Zukunft der Grammatik – Die Grammatik der Zukunft. Festschrift für Werner Abraham*. Tübingen: Stauffenburg.
- Christ, Rüdiger. 2014. *Zur Satzfügung im Deutschen*. Tübingen: Stauffenburg.
- Coniglio, Marco. 2011. *Die Syntax der deutschen Modalpartikeln: Ihre Distribution und Lizenzierung in Haupt- und Nebensätzen* (Studia Grammatica 73). Berlin: Akademie-Verlag. DOI: <https://doi.org/10.1524/9783050053578>
- Eberhardt, Ira. 2017. From a focus particle to a conjunction. German *zumal*: Diachronic and synchronic analysis. *Language*, section on *Historical Syntax, online version*. DOI: <https://doi.org/10.1353/lan.2017.0031>

- Featherston, Sam. 2005. Universals and grammaticality: Wh-constraints in German and English. *Linguistics* 43(4). 667–711. DOI: <https://doi.org/10.1515/ling.2005.43.4.667>
- Featherston, Sam. 2007. Data in generative grammar: The stick and the carrot. *Theoretical Linguistics* 33(3). 269–318. DOI: <https://doi.org/10.1515/TL.2007.020>
- Featherston, Sam. 2009. A scale for measuring well-formedness: Why linguistics needs boiling and freezing points. In Sam Featherston & Susanne Winkler (eds.), *The Fruits of Empirical Linguistics. Vol. 1 Process*, 47–74. Berlin: Mouton de Gruyter. DOI: <https://doi.org/10.1515/9783110216141.47>
- Featherston, Sam. 2019. The Decathlon Model. In András Kertész, Csilla Rákosi & Edith Moravcsik (eds.), *Current approaches to syntax – a comparative handbook*, 155–186. Berlin: Mouton de Gruyter. DOI: <https://doi.org/10.1515/9783110540253-006>
- Frey, Werner. 2011. Peripheral adverbial clauses, their licensing and the prefield in German. In Eva Breindl, Gisella Ferraresi & Anna Volodina (eds.), *Satzverknüpfung – Zur Interaktion von Form, Bedeutung und Diskursfunktion*, 41–77. Berlin: de Gruyter. DOI: <https://doi.org/10.1515/9783110252378.41>
- Gerbrich, Hannah, Vivian Schreier & Sam Featherston. 2019. Standard items for English judgement studies: Syntax and semantics. In Sam Featherston, Robin Hörnig, Sophie von Wietersheim, & Susanne Winkler (eds.), *Experiments in focus: Information structure and processing*, 305–328. Berlin: de Gruyter. DOI: <https://doi.org/10.1515/9783110623093-012>
- Grosz, Patrick. to appear. Discourse particles. In Lisa Matthewson, Cécile Meier, Hotze Rullmann & Thomas Ede Zimmermann (eds.), *The companion to semantics (SemCom)*. Oxford: Wiley.
- Günthner, Susanne. 1993. *Weil* – man kann es ja wissenschaftlich untersuchen – Diskurspragmatische Aspekte der Wortstellung in *Weil*-Sätzen. *Linguistische Berichte* 143. 73–115.
- Haegeman, Liliane. 2003. Conditional clauses: External and internal syntax. *Mind & Language* 18(4). 317–339. DOI: <https://doi.org/10.1111/1468-0017.00230>
- Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and the composition of the left periphery*. New York: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199858774.001.0001>
- Haiman, John & Sandra Thompson. 1984. ‘Subordination’ in Universal Grammar. In Claudia Brugman, Mónica Maccaulay, Amy Dahlstrom, Michelle Emanatian, Birch Moonwoman & Catherine O’Connor (eds.), *Proceedings of the tenth annual meeting of the Berkeley Linguistics Society*, 510–523. Berkeley, California: University of California Press. DOI: <https://doi.org/10.3765/bls.v10i0.1973>
- Keller, Frank. 2000. *Gradience in grammar: Experimental and computational aspects of degrees of grammaticality*. Edinburgh: University of Edinburgh dissertation.
- Küper, Christoph. 1991. Geht die Nebensatzstellung im Deutschen verloren? Zur pragmatischen Funktion der Wortstellung in Haupt- und Nebensätzen. *Deutsche Sprache* 19. 133–58.
- Neeleman, Ad & Reiko Vermeulen. 2011. Interpreting focus under negation. Ms, UCL/Ghent University.
- Pauly, Dennis. 2013. *Grenzfälle der Subordination: Merkmale, Empirie und Theorie abhängiger Nebensätze*. Potsdam: Potsdam University dissertation.
- Poulton, Eustace. 1989. *Bias in quantifying judgements*. Hove: Erlbaum.
- Rapp, Irene. 2018. Wenn man versucht, JA nichts Falsches zu sagen: zum Auftreten von Modalpartikeln in Haupt- und Nebensätzen. *Linguistische Berichte* 254. 183–228.

- Reis, Marga. 1997. Zum syntaktischen Status unselbständiger Verbzweit-Sätze. In Christa Dürscheid, Karl-Heinz Ramers & Monika Schwarz (eds.), *Sprache im Fokus. Festschrift für Heinz Vater zum 65. Geburtstag*, 121–144. Tübingen: Niemeyer.
- Reis, Marga. 2013. 'Weil-V2'-Sätze und (k)ein Ende? Anmerkungen zur Analyse von Antomo & Steinbach (2010). *Zeitschrift für Sprachwissenschaft* 32. 221–262. DOI: <https://doi.org/10.1515/zfs-2013-0008>
- Reis, Marga & Angelika Wöllstein. 2010. Zur Grammatik (vor allem) konditionaler V1-Gefüge im Deutschen. *Zeitschrift für Sprachwissenschaft* 29(1). 111–180. DOI: <https://doi.org/10.1515/zfsw.2010.004>
- Speyer, Augustin. 2011. Zur Integriertheit kausaler (Neben-)Sätze im Frühneu-hochdeutschen. *Sprachwissenschaft* 36. 53–84.
- Sweetser, Eve. 1990. *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511620904>
- Wegener, Heide. 1993. Weil – das hat schon seinen Grund. Zur Verbstellung in Kausalsätzen mit *weil* in gegenwärtigem Deutsch. *Deutsche Sprache* 21. 289–305.
- von Wietersheim, Sophie. 2016. Variable binding as evidence for clausal attachment. In Ingo Reich & Augustin Speyer (eds.), *Co- and subordination in German and other languages. Linguistische Berichte, Sonderheft* 21. 319–345. Hamburg: Buske.
- von Wietersheim, Sophie. in preparation. The syntactic integration of adverbial clauses: Experimental evidence from binding and other anaphoric relations. Tübingen: Tübingen University dissertation.
- von Wietersheim, Sophie & Sam Featherston. 2019. Does structural binding correlate with degrees of functional dependence? In Karin Beijering, Gunther Kaltenböck & Maria Sol Sansiñena (eds.), *Insubordination: Theoretical and empirical issues* (Trends in Linguistics Studies and Monographs 326), 265–290. Berlin: De Gruyter Mouton. DOI: <https://doi.org/10.1515/9783110638288-009>

How to cite this article: Eberhardt, Ira and Sam Featherston. 2019. Focus particles and negative scope: Both evidence for syntactic integration? *Glossa: a journal of general linguistics* 4(1): 126.1–18. DOI: <https://doi.org/10.5334/gjgl.613>

Submitted: 16 January 2018 **Accepted:** 23 July 2019 **Published:** 15 November 2019

Copyright: © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



Glossa: a journal of general linguistics is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS