The acquisition of the negative polarity item *any* in L2 English by L1 German speakers

TOM RANKIN
THOMAS WAGNER

*Author affiliations can be found in the back matter of this article

ABSTRACT

The study explores the acquisition of properties of the English existential quantifier *any* by German-speaking learners of English. The English existential quantifier is of particular theoretical interest, since its subtle grammatical constraints are often unobservable and thus may induce learnability difficulties, especially if parallel constructions in the L1 are scarce or entirely absent. In addition, the properties are underspecified in pedagogical materials, which renders them an apt test case for the exploration of input and taught knowledge in relation to acquisitional processes. The research question asks whether observable and/or learned grammar knowledge can shape acquired knowledge as measured by means of acceptability judgement tasks. To test this, 72 German-speaking C1 learners of English provided paced acceptability judgement ratings of sentences with any, which systematically differed according to grammatical and acquisitional constraints, partly replicating an experimental design by Marsden et al. (2018). Our results suggest sensitivity of German learners of English towards the grammatical constraints of *any*. This sensitivity seems to partly rest on explicit pedagogical input, with minimal L1 influence, but with some marked deviations from previous findings.

CORRESPONDING AUTHOR:

Tom Rankin

Masaryk University, CZ

tom.rankin@phil.muni.cz

1. INTRODUCTION

Gil, Marsden and Whong (henceforth GMW, where reference to a specific paper is not required) have explored the acquisition of existential quantifiers in a range of second languages by speakers of different first languages (Gil & Marsden, 2013; Gil et al., 2019; Marsden et al., 2018). The acquisition of paradigms such as English *any*, or Korean *wh*-existential *nwukwu* ('who/anyone/someone') is of theoretical interest for various reasons. First, their grammatical licensing and semantic interpretation are subject to subtle constraints. Patterns of acquisition thus furnish insight into how learners may overcome learnability issues at the syntax-semantics interface. Second, and linked to this, is the question of L1 influence in the acquisition of L2 grammars, given only partial overlap in the distribution and interpretation of related paradigms in different L1-L2 pairings. L2 learners' acquisition therefore furnishes insights into the nature of transfer and grammar restructuring in response to input patterns (see Gil & Marsden, 2013, based on Lardiere, 2009). And third, some properties of these items are taught in foreign language instruction, so that one may be able identify effects of explicit instruction in connection with, or in competition with, properties of acquisition.

In the present study, a paced acceptability judgment task from Marsden et al. (2018) was partly replicated. The replication complements the existing research by adding evidence from learners with L1 German. The syntactic and semantic properties of the relevant German features differ markedly from the L1s previously tested, resulting in further evidence relevant to potential L1 influence. The sentence types tested in the study reflected properties of the English existential quantifier *any* as a negative polarity item (NPI). These properties guided the research questions and hypotheses. We ask whether observable and/or learned grammar properties shape L2 knowledge, and in how far this L2 knowledge was potentially influenced by L1 transfer.

We proceed as follows. First, we will give an account of *any* as a negative polarity item and then summarise studies on the L2 acquisition of NPI *any*. Subsequently, we outline properties of related German existential quantifier paradigms. This serves as the basis for presentation of the experiment. Discussion of the results and suggestions of potential avenues for further research close the paper.

1.1 THE DISTRIBUTION OF NPI *ANY* IN ENGLISH

In English, *any* as an NPI requires a negative environment. Typically, this means occurring in the scope of sentential negation, as in (1). The occurrence of *any* outside the scope of negation is ungrammatical, as in (2).

(1)   I did not read any book.
(2)   *Anyone did not read the book.

Beyond this basic case, there are various subtleties regarding the precise characterisation of NPI licensing, and how semantic properties interact with the lexical requirements of NPIs, which are more or less stringent in the type of negative context required (see Giannakidou, 2011). To exemplify, *any* is licensed in questions, see (3). While questions do not involve overt negation, they can be analysed as related to nonveridicality. That is, involving the truth value of the content of the clause, or assertion of whether the event or state of affairs under consideration actually happened / will happen (see Giannakidou, 1998; Gil & Marsden, 2013).

(3)   Has anyone already read the book? (cf. *Anyone read the book.)

Any also occurs without overt negation in contexts where a nonveridical or negative inference can be computed (see Xiang et al, 2015). It is therefore licensed in clauses embedded under a semantically negative verb or within the scope of semantically negative adverbs, as in (4) and (5) respectively, from Gil et al. (2019, p. 221).

(4)   a.   John regrets that he ate anything at the party.
      b.   *John thinks that he ate anything at the party.

(5)   a.   John hardly ate anything at the party.
      b.   *John probably ate anything at the party.

In L2 English, acquisition involves two related issues. Firstly, the lexical representation of *any* must specify it as a negative polarity item, so that its distribution is restricted. In syntactic terms, one could propose that this means *any* bears an NPI feature (Gil et al, 2019), or an uninterpretable nonveridical feature (Gil & Marsden, 2013), which is checked in specific syntactic configurations. The second issue involves establishing what these configurations are, i.e. that an overt negative operator such as *not*, an interrogative operator, or an implicit negator introduced by the semantics of certain verbs and adverbs can create the requisite grammatical environments.

1.2 THE ACQUISITION OF *ANY* IN L2 ENGLISH

The use of *any* is often addressed in instruction in English as a foreign language (EFL, e.g., Berry, 2015). However, the information typically provided in instruction underdetermines the range of properties outlined above. This leads GMW to ask how L2 knowledge of a linguistic phenomenon develops when certain properties are taught, others are not taught but may be observable in incidental input, and still others are neither taught nor observable in input.

Based on studies of online EFL teaching materials (Marsden et al, 2018) and international EFL textbooks (Gil et al, 2019), it is shown that the typical pedagogical approach to *any* involves contrasting its use with *some* and characterising the choice as determined by clause type. *Any* occurs in questions and negative declaratives, while *some* occurs in affirmative declaratives.[1] Additional information may be provided as 'exceptions' to this basic rule, i.e., that certain adverbs collocate with *any*. Learners who have received such instruction would be able to correctly accept *any* with sentential negation and in questions (assuming such knowledge is accessible during performance). However, an overgeneralisation of this rule of thumb would complicate knowledge of more subtle properties. The absence of overt negation, as in (4a) and (5a) would lead learners to reject these sentences if they are relying on the 'textbook-rule' of using *any* with overt negation. Similarly, the presence of negation in a sentence like (2) could conceivably mislead learners into accepting this pattern due to the occurrence of a negator, even though the scope properties in this configuration rule it out.

Using paced acceptability judgement tasks, GMW investigated the acquisition of NPI properties of *any* by L1 Arabic and L1 Chinese-speaking learners. The aims of these studies were to explore whether learners' judgements indicate an ability to reject the less observable properties of *any*'s distribution dependent on scope and licensing from implicit negation, and whether learners might potentially overgeneralise textbook rules. In Marsden et al. (2018), the judgement task elicited acceptability ratings for 8 sentence types, reflecting the NPI properties and pedagogical rules discussed above, with both grammatical and ungrammatical stimuli. In their study, 4 tokens of each type resulted in 32 experimental items, plus 32 fillers. Table 1 summarises result for the eight sentence types.

All groups had lower accuracy on ungrammatical sentence types, although the results of *t*-tests on the grammatical/ungrammatical pairs were not significant for the L1 English group, indicating that they were equally able to accept grammatical clauses and reject ungrammatical ones. They also generally had higher accuracy across the board compared to learners. Overall, both learner groups performed similarly. Accuracy was highest on those contexts which tested the typically taught pedagogical rules: the difference between questions and declaratives and the fact that *any* occurs with the sentential negator *not*. Judgements are less consistent on those contexts which are not covered in pedagogical materials. However, this does not demonstrate that the pedagogical rule is overgeneralised. Such overgeneralisation would be evidenced by rejection of sentences that do not involve overt negation. However, learners can distinguish grammaticality based on the semantics of the licensing verb or adverb, even if judgements are markedly less robust than the taught cases. This leads GMW to suggest that instruction may have a facilitative effect for EFL learners.

For L1 Arabic and L1 Chinese learners of English, a facilitative effect would also be expected from L1 transfer. The corresponding Najdi Arabic *ʔayy* and Chinese *renhe/wh*-existentials have

---

[1] The key point is that where the issue is addressed in instruction, it is most likely in the form of a simplified rule, so that learners' explicit knowledge would not provide coverage of the full range of properties. The nature and use of pedagogical versus linguistic rules leads to additional questions, which are not discussed here (see for example Berry, 2021, Ch. 3; Swan, 1994).

SENTENCE TYPE	STATUS OF LICENSOR/ OPERATOR	GRAMMATI-CALITY	LEXICALISATION	SAMPLES		
				L1 ARABIC (*N* = 25)	L1 CHINESE (*N* = 22)	L1 ENGLISH (*N* = 15)
question	interrogative operator, non-veridical	grammatical	*Do you have any homework today?*	3.84	3.86	3.93
affirmative declarative	—	ungrammatical	**I've heard any news about the campaign*	3.08	2.86	3.73
negative declarative	negative operator	grammatical	*The teacher did not set any homework*	3.68	3.91	4.00
declarative outside scope	—	ungrammatical	**Anyone did not follow the instructions*	2.32	2.68	3.87
negative verb	implicit negation	grammatical	*I'm sorry I said anything about your driving test*	2.88	2.23	3.73
non-factive verb	—	ungrammatical	**I guess that you know anything about my visit*	2.12	2.09	3.60
negative adverb	implicit negation	grammatical	*James hardly ate anything at the party*	2.92	3.36	3.93
possibility adverb	—	ungrammatical	**Lucy probably bought anything last week*	2.52	2.41	3.73

Table 1 Sentence types for acceptability judgement tasks, adapted from Marsden et al. (2018), and mean accuracy of acceptability judgement tasks by sentence type, adapted from Gil et al. (2019).

a similar NPI distribution to *any*. The results provide limited grounds for assuming general positive transfer. There may be specific L1 effects for the L1 Chinese-speaking learners, who demonstrate a greater difference in acceptance of negative verbs versus negative adverbs as licensors (average accuracy 2.23 versus 3.36, respectively, on grammatical conditions). By contrast, the L1 Arabic group does not discriminate (2.88 grammatical negative verb versus 2.92 grammatical negative adverb). Gil et al. (2019, pp. 230–31) suggest that the form of Chinese negative adverbs may facilitate this distinction as the negator *bu* is often incorporated in these adverbs, e.g., *jihu bu – almost not* 'hardly'. The implicit negative meaning of such English adverbs is rendered more transparent for L1 Chinese learners given the association with an overt negator in the corresponding L1 lexical items.

A final relevant finding is that at least some learners could discriminate consistently between the different conditions and so seemed to have completely acquired the relevant linguistic properties of NPI *any* and its licensing conditions. Individual accuracy was calculated on the basis of rejecting 3 out of 4 tokens in each of the ungrammatical conditions and accepting 3 out of 4 in the grammatical conditions. On this measure, 9 of 22 L1 Chinese participants could consistently and accurately distinguish appropriate versus inappropriate syntactic contexts for *any*, and 10 of 25 advanced-proficiency L1 Arabic speakers.

1.3 PROPERTIES OF GERMAN

The similarities between English, Arabic and Chinese with respect to the distribution of existential NPI *any* may facilitate knowledge of NPI distribution. The overall patterns of judgements by L1 Arabic and L1 Chinese-speaking learners do not bear this out in detail. Learners demonstrate much less robust knowledge in environments involving lexical-semantic negation from verbs or adverbs licensing *any*. Given that Chinese and Arabic pattern similarly to English in pertinent respects, an interesting exploration of L1 effects would come from how learners acquire the distribution of *any* when the L1 provides no facilitative role. This is the case with German.

The closest equivalent to *any* is represented by forms with the indefinite marker *irgend-*. The *irgend*-series shares a number of functions with English *any* (see Haspelmath, 1997). This overlap motivates the assumption that the learner grammar will map *irgend*- onto *any* in a Feature Reassembly model (Lardiere, 2009). However, the syntax and semantics of the items differ in some subtle ways, potentially leading to learnability issues. Most pertinent as motivation for this study is the fact that the *irgend*-series is not subject to NPI distributional requirements. As we will see later in the discussion, clarifying the specific pattern of results

requires a more detailed study of specific interpretive properties. We return to this below as it is necessarily somewhat speculative given that this was not included as a variable in the replication study.

The *irgend*-series is formed by prefixing *irgend* to existential or indefinite phrases (*wh*-words, pronouns such as *jemand* someone or determiners such as *ein-* [=a]). Where there is optionality between a non-prefixed existential and an *irgend*-compound, *irgend-* introduces an implication of ignorance or indifference on the part of the speaker as in (6) (see Haspelmath, 1997, p. 45; Kratzer & Shimoyama, 2002).

(6) a.   *Hier habe ich was gefunden.*
          here have I what found
          'I have found something here.'

     b.   *Hier habe ich irgendwas gefunden.*
          here have I irgend.what found.
          'I have found something or other here'.

While *irgend*-forms are not NPIs (as evidenced by 6), syntactic context does affect their interpretation as free choice *any(thing)* versus non-specific *some(thing)* (see Penka, 2020, p. 653). Sentences 7–10, which translate the ungrammatical English sentences in Table 1, illustrate these interpretations. In each case, an *irgend*-form is possible, but will be interpreted as non-specific something/one rather than as an NPI.

(7) *Ich habe schon irgendwelche Neuigkeiten gehört.*   AFFIRMATIVE DECLARATIVE
     I have already irgend.which new things heard
     'I have heard some news or other.'
     (cf. *I have heard any news)

(8) *Irgendwer ist den Anleitungen nicht gefolgt.*   OUTSIDE NEG SCOPE
     irgend.who is the instructions not followed
     'Someone or other did not follow the instructions'
     (cf. *Anyone did not follow the instructions)

(9) *Ich nehme an, du weißt schon irgendwas über meinen Besuch.*   NON-FACTIVE V
     I take PRT you know already irgend.what over my visit
     'I guess that you've heard something or other about my visit.'
     (cf. *I guess that you know anything about my visit.)

(10) *Lucy hat wahrscheinlich letzte Woche irgendwas gekauft.*   POSSIBILITY ADV
     Lucy has probably last week irgend.what bought.
     'Lucy probably bought something or other last week.'
     (cf. *Lucy probably bought anything last week)

The *irgend*-series also occurs in the grammatical contexts for NPI *any*, illustrated in (11)–(15). A slight complication comes from the fact that German has a negative quantifier *kein-* which expresses *not-any*. An *irgend-* compound under the sentential negator is not possible without a specific pragmatic intention. For the sake of completeness, example (13) illustrates this pattern. The implication in this sentence is that the homework assignment is something special, i.e. "not just any old homework".

(11) *Hast du irgendwelche Hausaufagaben?*   QUESTION
     Have you irgend.which homeworks?
     'Do you have any homework (at all)?'

(12) *Der Lehrer hat keine Hausaufgaben erteilt.*   NEGATIVE DECLARATIVE
     The teacher has no homeworks distributed
     'The teacher didn't set any homework'

(13) *Die Lehrerin hat nicht irgendwelche Hausaufgaben erteilt.*   NEGATIVE DECLARATIVE
     the teacher has not irgend.which homeworks distributed.
     'The teacher didn't set just any old homework.'

(14) *Ich bedauere, daß ich irgendwas über deine Fahrprüfung gesagt habe.* NEGATIVE V
I regret that I irgend.what over your driving test said have.
'I'm sorry I said anything (at all) about your driving test.'

(15) *James hat bei der Feier kaum irgendwas gegessen.* NEGATIVE ADV
James has at the part hardly irgend.what eaten.
'James hardly ate anything at all at the party.'

Let's return to questions of learnability that this comparison raises. As already noted, Feature Reassembly (Lardiere, 2009) would assume that L1 German-L2 English learners map the lexical item *any* onto the features associated with *irgend-*, due to the shared functions and distribution. However, one element of the subsequent learning task is to restrict the distribution of *any* as an NPI.[2] In all contexts where *any* occurs, a representation which implicates the existential indefinite properties mapped from *irgend-* would grammatically license the occurrence, but potentially assign a different interpretation (see above). Crucially from a learnability perspective, ungrammatical usages would also be licensed by this representation, and on the basis of positive evidence, restructuring poses a problem. The presence of the *some/any* distinction in English might aid in arriving at a target representation. Nevertheless, without negative evidence, it is difficult to rule out ungrammatical occurrences of *any* in L1 German-L2 English. In many contexts, there is optionality between *some* and *any,* complicating the task of arriving at a target NPI representation which specifically disallows *any* in certain contexts. In a judgement task, where ungrammatical occurrences are presented, learners who have not reassembled features so as to restrict the distribution will syntactically license the usage (though potentially with different meanings).

Of course, as noted in the studies of teaching materials by GMW, negative evidence is available from instruction to the effect that *any* is restricted to overtly negative contexts or questions as compared to *some* (see Tesch, 1990 for similar findings from EFL materials for German-speaking learners). In common with GMW's predictions, the signature of reliance on a learned pedagogical rule would be a rejection of grammatical sentences where licensing of *any* is due to semantic features of verbs or adverbs, and acceptance of ungrammatical sentences with *any* outside the scope of negation.

Based on the discussion so far, and the insights furnished by GMW's experiments, we explore potential effects of pedagogical rules in explicit input, versus observability in incidental input. The hypothesised interplay between these factors and the eight sentence types is outlined in Table 2.

SENTENCE TYPE		GRAMMATICALITY	POTENTIAL SOURCES OF KNOWLEDGE	
			TAUGHT PEDAGOGICAL RULES IN EXPLICIT INPUT	NOT TAUGHT OBSERVABILITY IN INCIDENTAL INPUT
1	Questions	grammatical	+	+
2	Negative declarative	grammatical	+	+
3	Negative verb	grammatical	–	+
4	Negative adverb	grammatical	–	+
5	Affirmative declaratives	ungrammatical	+	–
6	Declarative outside scope	ungrammatical	–	–
7	Non-factive verb	ungrammatical	–	–
8	Possibility adverb	ungrammatical	–	–

Table 2 Interplay between potential sources of knowledge of *any* and the eight sentence types. *Note*: – = a particular source of knowledge is not associated with a sentence type. + = a particular source of knowledge is associated with a sentence type.

2   Gil & Marsden (2013) discuss learnability from the perspective of a particular syntactic analysis. They propose an uninterpretable nonveridical feature [uNV] to explain the distribution of existentials, and an additional feature [uNVα] for existentials with stronger distributional restrictions. Feature reassembly then involves selecting this feature for appropriate lexical items. We do not pursue this here as the experiment tests NPI contexts rather than nonveridicality as such.

As discussed in section 1.2, pedagogical rules focus strongly on occurrences in questions and negative declaratives, facilitating knowledge of sentence types 1, 2, and – as a common pedagogical counter example – sentence type 5. Sentences involving *any* licensed by semantically negative verbs and adverbs occur in incidental input. Knowledge of these is therefore in principle derivable from positive evidence, even though these features are not typically taught. Ungrammatical occurrences of types 6–8 in Table 2 are not taught. They are by definition not available from incidental input as ungrammatical sentences do not occur, and they would not be ruled out by mapping from L1 German.

Key comparisons are therefore (i) between the taught grammatical questions and negative declaratives versus untaught types with negative verb and adverbs. A difference in the acceptability ratings would index potential effects of instruction. Such instructional effects can be further explored by (ii) comparing the rating accuracy of grammatical questions (1) versus parallel ungrammatical affirmative declaratives (5), and also by contrasting negative declaratives (2) with sentences containing scope violations (6). Because the ungrammatical sentences are by definition absent from positive input, teaching effects would be indexed by more accurate ratings of sentence type 1 versus 5 and 2 versus 6.

Finally, input and L1 effects would be identified by comparisons of (iii) grammatical, observable but untaught sentence types with negative verbs (3) and adverbs (4) versus ungrammatical ones with non-factive verbs (7) and possibility adverbs (8), which are neither taught nor observable. A tendency to reject ungrammatical sentences altogether would indicate an overgeneralisation of the pedagogical rule, while a tendency to accept all of them would indicate L1 effects. More accurate ratings in sentence types 3 and 4 against 7 and 8 would furthermore point towards developing grammatical L2 knowledge despite the absence of an overt licensor.

2. THE STUDY

2.1 PARTICIPANTS

72 learners of English studying in an English Language Teaching degree programme participated in the study. This sample size would permit detection of effects as small as $\eta_p^2 = 0.13$ (Kumle, Võ & Draschkow, 2021; Schütze, 2016). Participants' mean age was 22.48 years, (*SD* = 3.29), and German was their L1, either mono- or multilingually. 53 identified as female, 18 as male, and one as non-binary. All participants had had eight years of regular, instructed EFL school teaching prior to their tertiary studies; in addition to that, around 11% reported a stay abroad between one and six months, while around 14% had stayed for a longer period of time in an English-speaking country. The great majority, though, had never had an extended stay in an English-speaking environment (75%).

All 72 participants took part voluntarily, anonymously, and with explicit consent, but without any financial renumeration. At the time of testing, participants had been studying English at tertiary level for at least two years and were approaching C1 proficiency level (CoE, 2001), i.e. "advanced" as defined by the Common European Framework of reference for language. This was measured by proficiency tests at the beginning of the degree and after two years in the study programme. The data of all 72 participants were used for reliability and bias analyses of the acceptability judgement ratings.

2.2 MATERIALS AND METHOD

As in the GMW studies, participants rated the naturalness of sentences containing *any* in a paced acceptability judgement task. 12 lexicalisations for each of the sentence types were created, resulting in 96 tokens, which were distributed in pseudo-randomised experimental lists following a Latin Square design. Participants saw 24 randomised test items (3 items per experimental condition) as well as eight fillers, either systematic distractors or items piloting a separate study on negative inversion. Test items, raw data and R-script are accessible via https://osf.io/9qt46/.

The acceptability judgement task was administered online through SoSciSurvey (Leiner, 2021). Participants were informed that the experiment was part of a study about learners' grammatical knowledge, the procedure was explained briefly, and two untimed items for practice were provided.

The sentences were rated on a four-point Likert scale, labelled *I'm sure this sounds natural, I think this sounds natural, I think this does not sound natural,* and *I'm sure this does not sound natural.* Since the acceptability judgement scores were intended to elicit participants' intuition about the naturalness of a given sentence, a gradable rating option was chosen instead of a binary right-wrong choice. Contrary to Marsden et al. (2018), an even point-scale was chosen so that participants were forced into one or the other direction, thereby avoiding clusters in the classic middle-ground option *cannot decide.* For the same reason, the option *I do not know* was omitted. Thus, it was assumed that even mild and subtle intuitions about a sentence's acceptability could be captured.

2.3 DATA CODING AND ANALYSIS

Signal Detection Theory (SDT) was used to analyse how participants discriminated between grammatical and ungrammatical sentence types and whether they showed any general bias towards accepting or rejecting (Huang & Ferreira, 2020). Huang & Ferreira (2020) propose using SDT for judgement data as it delivers insights into how well participants differentiate between grammatical and ungrammatical sentences and whether they have any bias in their preferences. To this end, acceptability ratings were coded as follows. Detecting a grammatical sentence's grammaticality resulted in a *hit*; rejecting a grammatical sentence resulted in a *miss.* If participants correctly detected the ungrammaticality of a sentence, they scored a *correct rejection,* or *true negative.* Where they rated ungrammatical sentences as grammatical, they scored a *false alarm.* Based on this 2×2 matrix, the SDT indices *d'* and *c* were calculated; *d'*, a measure of sensitivity, captures participants' reliability when distinguishing between grammatical and ungrammatical sentences. Index *c*, the criterion location index, expresses conservative or liberal biases.

Mixed ordinal regression modelled potential effects of sentence type on the acceptability ratings. The four Likert points were treatment-coded as ordered categorical data, with the levels "1", "2", "3", and "4", "1" reflecting the least appropriate, and "4" reflecting the most appropriate rating. Thus, a rating such as *I'm sure this sounds natural* for a grammatical sentence would be coded as "4", most appropriate, while the same rating for an ungrammatical sentence would score a "1", least appropriate. The sentences were sum-contrast-coded, so that the predictor levels represent the difference between these levels and the grand mean, thus effectively 'centring' the effects of the variable's levels at this value (Schad et al., 2020). This coding was chosen since none of the eight sentence types could reasonably be thought of as a baseline level against which the other estimates would then be calculated.

For the mixed modelling, we used cumulative link mixed regression with Laplace approximation from the *ordinal* package (Christensen, 2019, version 2019.12.10) in R (R Core Team, 2021). Main regression effects from the cumulative link mixed model are reported based on Type-II Wald χ^2-tests, as implemented in the *RVAideMemoire* package (Hervé, 2022). Partial effects statistics (Wald) were derived from the model summaries produced within *ordinal.*

3. RESULTS

After replacing missing values by appropriate imputations, participants' rating accuracy in both the grammatical and ungrammatical sentences were quantified. The heavily left-skewed (Fisher-Pearson $g^1 = -1.30$) overall distribution towards category "4" (most appropriate) could be interpreted as a first indication of above-average accuracy. In order to examine if accuracy holds across grammatical and ungrammatical sentence types, the ratings were coded as illustrated in Table 3.

		GRAMMATICALITY	
		ACCEPTABLE	UNACCEPTABLE
ratings	acceptable	693 *hits*	134 *false alarms*
	unacceptable	150 *misses*	692 *correct rejections*

Table 3 Contingency table for the counts of the four SDT coding and resulting counts.

Figure 1 Conditional means plot for acceptability judgement ratings by sentence type.
Note: The dashed horizontal line represents the grand mean. The vertical arrows illustrate the distance between each mean sentence type rating and this grand mean. Error bars illustrate 95% confidence intervals.
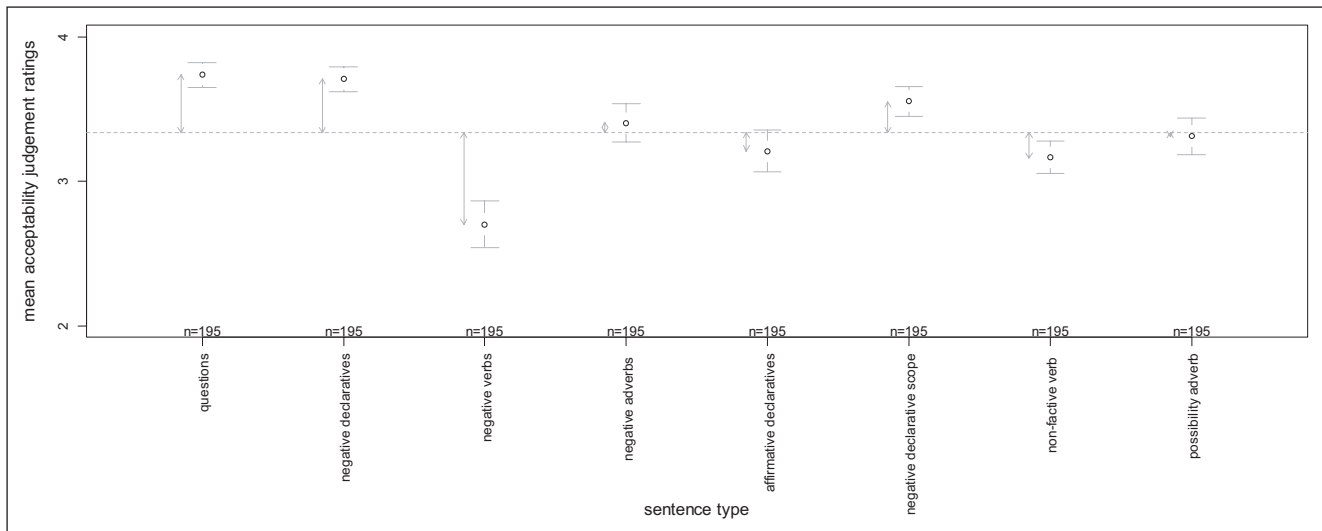
Table 3 is almost perfectly symmetrical, with about the same number of true positives (hits and correct rejections). This is a first indication of participants' appropriate sensitivity towards grammaticality. Accordingly, a *d'* value of 1.91 (*SE* = 0.01, 95% *CI* [1.90, 1.92], significant on an inter-individual level (*t*(71) = 22.42, *p* < .001), shows that the acceptability ratings are 40% away from chance level, which in turn suggests a rather reliable discrimination between grammatical and ungrammatical sentence types. In line with these results, there was hardly any bias towards one particular type of response (acceptable, unacceptable) in the participants' ratings either (*c*-value = .03, *t*(71) = 0.99, *p* = 0.32); in other words, the learners were neither particularly conservative nor liberal in their acceptability ratings.

Consider Figure 1; it visualises the conditional means for the participants' acceptability judgement ratings across the eight sentence types.

Table 4 Coefficients from a cumulative link mixed model (Laplace approximation) with random intercepts and slopes for participants.
Note: Number of partial effects = 8, number of observations = 1560, total *N* = 65, *CI* = confidence interval, *LL* = lower limit, *UL* = upper limit, OR = odds ratios, pseudo-*R²* (McFadden) = 0.09, Cox and Snell (ML) = 0.17, Cox & Snell's *R* squared (Cox & Snell, 1989) = 0.19. [a] Cohen's *d* effect size approximations were calculated using (*log*(OR) x√3)/π, Sánchez-Meca et al., 2003. Condition number of the Hessian, measuring the empirical identifiability of the model, is 0.02. Maximum absolute gradient of the log-likelihood function with respect to the parameters is 0.34.

EFFECTS	ESTIMATE		COHEN'S D APPROX.[a]	SE	Z	95% CI ODDS RATIOS		P
	LOGITS	ODDS RATIOS				LL	UL	
fixed effects								
threshold coefficients								
1 | 2	–3.26	0.04	–1.80	0.15	–22.23	0.03	0.05	<.001***
2 | 3	–2.12	0.12	–1.17	0.12	–17.64	0.09	0.15	<.001***
3 | 4	–0.48	0.62	–0.26	0.1	–4.67	0.51	0.76	<.001***
predictors								
questions	1.25	3.49	0.69	0.24	5.11	2.16	5.64	<.001***
negative declaratives	0.96	2.61	0.53	0.22	4.43	1.71	4	<.001***
negative verbs	–1.52	0.22	–0.84	0.2	–7.48	0.15	0.33	<.001***
negative adverbs	0.1	1.1	0.05	0.19	0.52	0.76	1.59	.60
affirm. declaratives	–0.31	0.73	–0.17	0.2	–1.58	0.5	1.08	.11
negative declarative	0.42	1.52	0.23	0.19	2.2	1.05	2.21	<.05*
non-factive verb	–0.69	0.5	–0.38	0.16	–4.24	0.37	0.69	<.001***
possibility adverb	–0.21	0.81	–0.12	0.17	–1.21	0.54	1.22	0.29
random effects	variance	SD	corr.					
intercept	0.94	0.97						
negative declaratives	0.47	0.69	–0.52					
negative verbs	1.74	1.32	–0.60	0.64				
negative adverbs	0.41	0.64	–0.49	0.87	0.60			
affirm. declaratives	1.66	1.29	–0.40	–0.18	–0.30	0.41		
negative declarative	1.71	1.31	–0.73	0.41	0.48	0.10	0.66	
non-factive verb	1.54	1.24	–0.55	0.20	0.12	–0.15	0.87	0.93
possibility adverb	1.23	1.11	–0.44	0.08	–0.05	–0.27	0.92	0.85	0.99

There are pronounced differences in conditional means; sentence types 1, 2, 4, and 6 elicited ratings above the grand mean, while others are below, with sentence type 3 deviating most markedly. The conditional means across all eight sentence types are significant (Type II Wald ANOVA, $\chi^2(7)$ = 91.77, p < .001), and differences in individual repeated measure ratings were modelled using a cumulative link mixed model with Laplace approximation. Random intercepts and slopes for participants were included, as their variance was substantial (S^2 = 0.21, SD = 0.47); item variance, in contrast, was minute (values < .001). The model's fixed and random effects are illustrated in Table 4.

Table 4 shows that five out of the eight sentence types significantly influence the acceptability ratings. Medium to large effect sizes ($d \geq 0.5$, Cohen, 1988), however, can be seen for sentence types 1, 2, and 3 only. Here, questions and declaratives show odds ratios greater than 1, while negative verbs show a ratio below 1. The odds ratio of 0.22 for sentences containing negative verbs, for instance, means that the odds of getting a more accurate acceptability judgement rating is 0.22 times that of the grand mean, or 78% lower than with sentences on average, holding all other sentence types constant; in other words, this model would predict that participants rate this sentence type much less accurately than they do on average. While this model suffers from minor scale effects, as the proportional odds assumption does not hold perfectly across all predictor levels, the estimates still provide a useful and reliable measurement of the sentence type effects (Harrell, 2015, p. 316; Peterson & Harrell, 1990).

So far, analyses have shown that participants can reliably, and in an unbiased fashion, discriminate between grammatical and ungrammatical sentences. Also, the accuracy of participants' ratings clearly depends on licensing conditions for *any*; some sentence types trigger better, some worse ratings than on average. Particular difficulties seem to arise with sentences containing negative verbs. This is noteworthy: learners are more reliably able to reject ungrammatical *any* embedded under non-factive verbs than they are able to accept grammatical occurrences licensed by negative verbs. Indeed, grammatical use of *any* licensed by negative verbs produces the worst ratings of all sentence types.

We then examined the influence of potential sources of knowledge, as outlined in Table 2. Recall that this comprised the five sentence-pair and pooled sentence type contrasts 1+2 versus 3+4, 1 versus 5, 2 versus 6, 3+4 versus 7+8, and 1–4 versus 5–8. In addition, the contrast between sentence type 3 and 7 was explored, comparing grammatical negative verbs with ungrammatical non-factive verbs. Table 5 and Figure 2 illustrate the effects for each of these contrasts, based on monofactorial cumulative link mixed models with random intercepts and slopes for participants. Significant pairs in Table 5 are also mirrored in estimated marginal means contrasts (Tukey-adjusted *p*-values).

		MAIN EFFECTS	PARTIAL EFFECTS			
		ANOVA WALD STATISTICS	ESTIMATE	*SE*	*Z*	*P*
1+2 vs 3+4	taught – untaught	$\chi^2(1)$ = 65.97	0.81	0.09	8.87	<.001***
1 vs 5	questions – affirmative decl.	$\chi^2(1)$ = 26.42	0.75	0.15	4.90	<.001***
3 vs 7	neg. verbs – non-factive verbs	$\chi^2(1)$ = 8.47	−0.44	0.15	−2.95	<.01**
2 vs 6	neg. decl. – neg. decl. scope	$\chi^2(1)$ = 2.84	0.26	0.16	1.67	.09
3+4 vs 7+8	licensor – no licensor	$\chi^2(1)$ = 1.34	0.25	0.22	1.16	.25
1–4 vs 5–8	grammatical – ungrammatical	$\chi^2(1)$ = 2.08	0.13	0.09	1.47	.14

Table 5 Main and partial fixed effects for six sentence pairs and pooled sentence contrasts.

As shown in Table 5 and Figure 2, significant contrasts arise between taught and untaught sentence types, questions versus affirmative declaratives, and sentences containing negative versus non-factive verbs. In the top-left panel of Figure 2, the significant positive slope towards taught sentences shows that participants produced significantly more accurate acceptability judgement ratings for sentences representing typically taught properties compared to the untaught properties. In the top middle panel, we can see that questions elicited significantly more accurate ratings than affirmative declaratives, which might indicate an effect of instruction given that questions are taught as a prototypical grammatical context for the use of *any*. The difference between sentences involving negative and non-factive verbs, visible

Figure 2 Conditional means plots for six sentence pairs and pooled sentence contrasts.
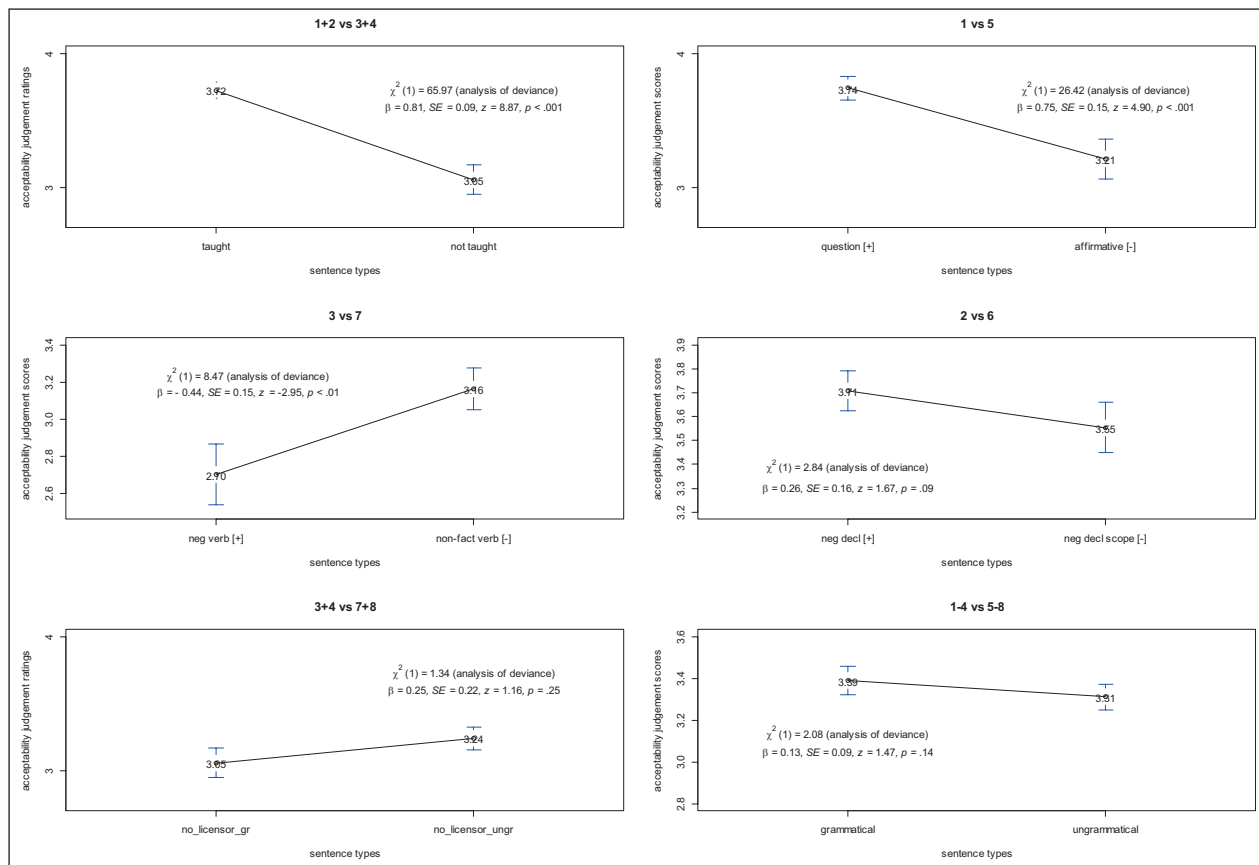
in the top-right panel, is, again, surprising. While this property is not generally addressed in instruction, learners can encounter the licensing of *any* from negative verbs in positive input. However, this type of grammatical sentence received a significantly less accurate rating than the ungrammatical condition with non-factive verbs. Such an effect cannot be observed for the two adverb conditions (sentence types 4 and 8), indicating that learners were reliably able to accept grammatical sentences and reject ungrammatical counterparts in the face of a similar acquisition task as that involving verbs. As we can see in the bottom panels in Figure 2, scope, the absence of a licensor in grammatical and ungrammatical sentences as well as grammaticality in general do not significantly influence rating accuracy.

4. DISCUSSION AND CONCLUSIONS

Overall, the results indicate an ability to distinguish acceptability of NPI contexts for the occurrence of *any* by advanced L1 German-speaking learners of English, with a complication related to judgements of licensing from verb semantics. The discussion considers the potential sources of knowledge and interaction between them in leading to this pattern.

It can be stated that, by this level of proficiency, learner knowledge is not generally conditioned by L1 properties. Recall that the German *irgend*-series would be mapped to *any* in a Feature Reassembly framework and that this mapping would license the occurrence of *any* in all of the sentence types tested, meaning that L1 influence would be indexed by a tendency to accept ungrammatical sentences. Clearly, the learners in the study have restructured grammars that permit broad discrimination of licensing conditions. A question is the relative contribution of pedagogical and incidental input to this restructuring. Taking pedagogical input first, it seems clear that explicit pedagogical rules can have a facilitative effect, supporting GMW's results. Those licensing contexts typically addressed in pedagogical grammar evince significantly more consistent target judgements than conditions which are not addressed in instruction.

The other side of this coin is the question of untaught and (un)observable properties. Even if performance on the untaught properties is less robust, effects of observability would be evidenced by an ability to distinguish grammaticality on the basis of the semantics of licensing verbs or adverbs. In this area, the results from the present study diverge from the previous findings, as illustrated in Table 6.

			STUDIES			PRESENT STUDY
			GMW			
			L1 ARABIC (*N* = 25)	L1 CHINESE (*N* = 22)	L1 ENGLISH (*N* = 15)	L1 GERMAN (*N* = 65)
Sentence types	Q / Decl	Grammatical	3.84	3.86	3.93	3.74
		Ungrammatical	3.08	2.86	3.73	3.21
	Neg Scope	Grammatical	3.68	3.91	4.00	3.71
		Ungrammatical	2.32	2.68	3.87	3.55
	Main Verb	Grammatical	2.88	2.23	3.73	**2.70**
		Ungrammatical	2.12	2.09	3.60	**3.16**
	Adverb	Grammatical	2.92	3.36	3.93	3.41
		Ungrammatical	2.52	2.41	3.73	3.31

Table 6 Mean acceptability judgement ratings in all eight sentence types across four experimental studies.

Unlike previous studies, our learners produced considerably more accurate ratings for ungrammatical *any* embedded under non-factive verbs. As well as diverging from previous learner groups' results, this stands out as distinct from the native speakers in GMW's studies. For all these groups, grammatical sentences consistently evince more accurate judgements than ungrammatical sentences, but this pattern is reversed for the L1 German group's performance on the verb-licensing condition. Focussing only on the learners, performance on grammatical sentences with negative verbs is numerically similar (2.88 L1 Chinese, 2.70 L1 German, 2.23 L1 Arabic). The outlier that calls for explanation is specifically the better-than-expected performance in rejecting ungrammatical sentences. This result is puzzling and is not conducive to a straightforward analysis based on the learnability issues and L1 properties discussed above.[3]

We finish by offering some speculation on a possible analysis. This assumes the specific pattern of results is due to a confluence of factors involving complications in the interpretation of *any* licensed by verbs, compounded by L1 effects. This points to avenues for future research. It can be noted that while the German equivalents of the English main-verb conditions are *grammatically* possible, there is a subtle pragmatic/semantic distinction, which impinges upon the acceptability of the existential with the different types of matrix verbs. This is illustrated in (16) and (17), which extend example (4) from earlier.

(16) Maria bedauert, dass sie irgendwas bei der Feier gegessen hat.
      Maria regrets that she irgend.what at the party eaten has.
      'Maria regrets that she ate anything/something at the party.'

(17) Maria glaubt, dass sie irgendwas bei der Feier gegessen hat.
      Maria thinks that she irgend.what at the party eaten has.
      'Maria thinks that she ate something at the party.'

In (16), the *irgend*-existential can be interpreted as meaning that Maria did in fact eat something, but regrets it for example because she got food poisoning or was intending to diet that day, etc. Or it can mean that she regrets eating anything at all. These readings are pragmatically unexceptional and available for a sentence in isolation. By contrast, the most natural reading of (17), as indicated in the translation is as a non-specific *some*-existential, with the interpretation that Maria is unsure whether or not she actually ate some foodstuff. Remember that *irgend*- also functions to introduce an implication of uncertainty or ignorance. Such a reading for a sentence in isolation is obviously somewhat pragmatically odd. If there is in fact some level of L1 influence, resulting in this pragmatically odd reading, this may explain the more reliable rejection of the ungrammatical English sentences. In other words, the learners have acquired the basic NPI distributional patterns, but may still be influenced by L1 properties in the interpretation of existential *any* in other contexts, giving rise to free-choice

---

3   GMW mention a potential methodological issue specifically in the verb conditions. These involve bi-clausal sentences, unlike the other sentence types. This greater complexity may explain the generally lower accuracy across the groups on the verb conditions due to greater processing demands in the paced judgement task. However, it cannot explain the reversal of the grammatical/ungrammatical pattern specifically for the L1 German group.

or indefinite meanings. Of course, this raises the question of why the verb conditions pattern differently to the adverb conditions, which are also not the core NPI contexts. Grammatical adverb conditions would receive an NPI reading in German while ungrammatical conditions would also receive a non-specific *some* reading (see 18 and 19). However, in these cases, there are no additional complications with respect to pragmatic interpretation. So, if the idea that there is subtle L1 influence is on the right lines, we would not necessarily expect a difference to emerge in these conditions.

(18) Maria hat kaum irgendwas bei der Feier gegessen.
      Maria has hardly irgend.what at the party eaten
      'Maria hardly ate anything at the party.'

(19) Maria hat wahrscheinlich irgendwas bei der Feier gegessen.
      Maria has probably irgend.what at the party eaten.
      'Maria probably ate something or other at the party.'

Furthermore, judgements involving adverbs are generally better than the verb conditions across the board (see Table 6). It may be that there is something about adverbs that facilitates acquisition of NPI licensing prior to licensing from verbs. This something might be frequency if the adverb contexts occur more frequently in input.[4] It might also be another indirect pedagogical effect. Recall that GMW's survey of teaching materials found that pedagogical rules did sometimes mention co-occurrence of *any* with specific adverbs, even if this was often presented as an 'exception' to the rule. Realistically, it may be a confluence of this range of factors which leads to the differences.

The GMW studies, and by extension this replication, only focussed on NPI distribution of *any*. It is clear from the results from L1 German speakers learning English that future research should explore comprehension and pragmatics in order to gain a fuller picture of how *any* is acquired and whether NPI and indefinite or free-choice meanings and pragmatics are implicated. Especially for L1 German-speaking learners, results from acceptability may mask a more complex picture related to the semantic and pragmatic readings assigned to different types of sentences, which may implicate continued L1 influence on existentials and indefinites at the level of semantic and pragmatic interpretation.

AUTHOR AFFILIATIONS

Tom Rankin ⬤ orcid.org/0000-0001-5372-8773
Masaryk University, CZ

Thomas Wagner ⬤ orcid.org/0000-0003-4028-2527
University College of Education Upper Austria, AT

REFERENCES

Berry, R. (2015). Grammar myths. *Language Awareness, 24*(1), 15–37. DOI: https://doi.org/10.1080/09658 416.2013.873803

Berry, R. (2021). *Doing English grammar: Theory, description and practice*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/9781108325745

Christensen, R. H. B. (2019). *Ordinal. Regression models for ordinal data*. R package (Version 2019.12–10) [Computer software]. https://CRAN.R-project.org/package=ordinal

CoE (Council of Europe). (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Cohen, J. C. (1988). *Statistical power analysis for the behavioral sciences*. New York: Routledge. DOI: https://doi.org/10.4324/9780203771587

Cox, D. R., & **Snell, E. J.** (1989). *Analysis of binary data. Monographs on statistics and applied probability*. 2nd edition. London: Chapman and Hall.

Giannakidou, A. (1998). *Polarity sensitivity as (non)veridical dependency*. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/la.23

---

4   An anonymous reviewer raises the question of whether the occurrence of formulaic expressions such as *hardly anything* might facilitate acquisition of licensing from adverbs, thus predicting that performance might be lexically constrained (i.e. better on *hardly* as compared to *seldom*). This is an interesting avenue for future research.

Giannakidou, A. (2011). Negative and positive polarity items. In K. Heusinger, C. Maienborn & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning. Volume 2.* (pp. 1660–1712). Berlin, Boston: De Gruyter Mouton.

Gil, K.-H., & Marsden, H. (2013). Existential quantifiers in second language acquisition: A feature reassembly account. *Linguistic Approaches to Bilingualism*, *3*(2), 117–149. DOI: https://doi.org/10.1075/lab.3.2.01gil

Gil, K.-H., Marsden, H., & Whong, M. (2019). The meaning of negation in the second language classroom: Evidence from 'any'. *Language Teaching Research*, *23*(2), 218–236. DOI: https://doi.org/10.1177/1362168817740144

Harrell, F. E. (2015). Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis. 2nd edition. *Springer Series in Statistics*. Heidelberg et al.: Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-19425-7

Haspelmath, M. (1997). *Indefinite pronouns*. Oxford University Press.

Hervé, M. (2022). *RVAideMemoire: Testing and plotting procedures for biostatistics*. R package (Version 0.9-81-2) [Computer software]. http://127.0.0.1:22992/library/RVAideMemoire/html/RVAideMemoire-package.html

Huang, Y., & Ferreira, F. (2020). The application of signal detection theory to acceptability judgments. *Frontiers in Psychology, 11*(73). DOI: https://doi.org/10.3389/fpsyg.2020.00073

Kratzer, A., & Shimoyama, J. (2002). Indeterminate pronouns: The view from Japanese. In Y. Otsu (Ed.), *The proceedings of the Third Tokyo Conference on Psycholinguistics*, (pp. 1–25). Tokyo: Hituzi Syobo.

Kumle, L., Võ, M. L H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research Methods, 53*(6), 2528–2543. DOI: https://doi.org/10.3758/s13428-021-01546-0

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research*, *25*(2), 173–227. DOI: https://doi.org/10.1177/0267658308100283

Leiner, D. J. (2021). SoSci Survey (Version 3.2.31) [Computer software]. https://www.soscisurvey.de

Marsden, H., Whong, M., & Gil, K.-H. (2018). What's in the textbook and what's in the mind: Polarity item "any" in learner English. *Studies in Second Language Acquisition, 40*(1), 91–118. DOI: https://doi.org/10.1017/S0272263117000018

Penka, D. (2020). Negative and positive polarity items. In M. Putnam & B. Page (Eds.), *The Cambridge handbook of Germanic linguistics* (pp. 639–660). Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/9781108378291.028

Peterson, B., & Harrell, F. E. (1990). Partial proportional odds models for ordinal response variables. *Applied Statistics, 39*(2), 205–217. DOI: https://doi.org/10.2307/2347760

R Core Team. (2021). R: A language and environment for statistical computing (Version 4.1.2 "Bird Hippie") [Computer software]. Vienna, Austria: R-Foundation for Statistical Computing. Available at https://www.R-project.org/

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, *8*(4), 448–467. DOI: https://doi.org/10.1037/1082-989X.8.4.448

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language, 110*, 104038. DOI: https://doi.org/10.1016/j.jml.2019.104038

Schütze, C. T. (2016). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Berlin: Language Science Press. DOI: https://doi.org/10.26530/OAPEN_603356

Swan, M. (1994). Design criteria for pedagogic language rules. In M. Bygate, A. Tonkyn & E. Williams (Eds.), *Grammar and the language teacher* (pp. 45–55). New York: Prentice Hall.

Tesch, F. (1990). *Die Indefinitpronomina* some *und* any *im authentischen englischen Sprachgebrauch und in Lehrwerken: eine empirische Untersuchung*. [The indefinite pronouns some and any in authentic English usage and in teaching materials: An empirical exploration]. Tübingen: Narr Verlag.

Xiang, M., Grove, J., & Giannakidou, A. (2015). Semantic and pragmatic processes in the comprehension of negation: An event related potential study of negative polarity sensitivity. *Journal of Neurolinguistics, 38*, 71–88. DOI: https://doi.org/10.1016/j.jneuroling.2015.11.001

WHITE ROSE
UNIVERSITY PRESS
Universities of Leeds, Sheffield & York