

Does intonation automatically strengthen scalar implicatures?*

John Michael Tomlinson, Jr.
*Leibniz Center for General
Linguistics*

Camilo R. Ronderos
University of Oslo

Submitted 2020-10-05 / First decision 2021-03-01 / Revision received 2021-04-02 /
Second decision 2021-04-08 / Revision received 2021-04-13 / Accepted 2021-04-20 /
Published 2021-05-05 / Final typesetting 2023-02-28

Abstract Two mouse-tracking experiments tested predictions from two different models of scalar implicature as to whether exhaustive interpretations are computed prior to ignorance implicatures. We use different German intonational patterns to probe the availability of these interpretations (Experiments 1 and 2) and add a speaker competence manipulation in Experiment 2. Results from Experiment 1 found that deriving exhaustive interpretations with an L+H* was delayed to ignorance implicatures with an L*+H contour. Experiment 2 replicated this finding even with a strengthened competence assumption about the speaker. We interpret our processing data as providing constraints on the computational mechanisms underlying the interpretation of scalar implicatures.

Keywords: scalar implicature, intonation, competence assumption, mouse-tracking, prosody

* We thank Bob van Tiel, Guillermo del Pinal, and Nicole Gotzner for comments on prior drafts of the manuscripts. We would also like to thank two reviewers for their constructive feedback on earlier versions, and Kim Jördens for assisting with the data collection. The first author drafted and edited the manuscript and analyzed the mouse-tracking data. The second author helped edit versions of the manuscript and analyzed the forced-choice data. Both authors designed and implemented the experiments. The first author was supported by funding from the Alexander von Humboldt Foundation and the German Ministry of Education and Research (BMBF)

1 Introduction

Implicatures rely on integrating both linguistic and speaker-specific information (Grice 1975). Consider the following examples:

- (1) A: Were Manu and Moni at the party?
B: Manu was there.

While B could have addressed the attendance of Manu and Moni separately, B only confirms the presence of Manu. Here, A can interpret B's utterance in two ways. First, A might take B's utterance to mean that B lacks knowledge about Moni's attendance. On the other hand, B might assume that by *only* mentioning Manu, A should interpret B's answers in (1) *exhaustively* (Moni was *not* at the party). Traditionally such exhaustive interpretations of answers are thought to come about from *scalar implicatures* (from now on SIs): B used a weaker expression instead of a stronger one (*Only Manu was there*), effectively violating Grice 1975's Maxim of Quantity (Sauerland 2004, van Rooij & Schulz 2004, Spector 2007).¹

This kind of Gricean reasoning not only explains how speaker knowledge can lead to the derivation of SIs. It is also a useful framework for showing when and how SIs are not derived (Grice 1975). Consider the following example:

- (2) A: Hey, I heard that you could only briefly stop by Tom's party. Were Manu and Moni there?
B: Manu was there.

In this example, A has reason to doubt that B is well-informed or competent enough to even assert the stronger alternative. This is known as the *competence assumption* (van Rooij & Schulz 2004, Chemla 2008). In these cases, A should not interpret B's response in (2) exhaustively. Instead, they should derive an *ignorance implicature* (A is certain that B does not have enough

¹ Example (1) is not your run of the mill SI. Unlike the more canonical examples of SIs such as (<some, all>), (<and, or>) or various adjectival scales, e.g., (<attractive, good-looking>), the scalar alternatives (<Manu, Moni>) are not lexically pre-determined. Rather, the scales in (1) are determined by context, hence are known as *ad-hoc scalar implicatures* (Hirschberg 1985, Frank & Goodman 2012, Breheny, Ferguson & Katsos 2013). Although ad-hoc scales might seem less uniform than lexically determined scales, lexical scales also display variable or non-uniform behavior (Van Tiel et al. 2014). Our assumption is that all SIs share similar computational mechanisms. This claim is supported in a priming study by Bott & Chemla (2016) who found that quantifier and numerical SIs reliably prime ad-hoc SIs and vice versa.

Does intonation automatically strengthen scalar implicatures?

knowledge to assert the stronger alternative). The Gricean model posits an additional mechanism that allows for minimally enriched *weak implicatures* to be further strengthened to a (strong) SI, i.e., exhaustive interpretation, or weakened to an ignorance implicature (Sauerland 2004, Chemla & Singh 2014, Dieuleveut, Chemla & Spector 2019):

- (3) *Weak/primary implicatures*
It is not the case that Speaker B believes (Moni & Manu)
- (4) *Strong/secondary implicatures*
Speaker B believes that (\neg Moni, Manu)
- (5) *Ignorance implicatures*
Speaker B believes that (?Moni, Manu)

This process, known as the *epistemic step* (Sauerland 2004, Breheny, Ferguson & Katsos 2013), constitutes an additional reasoning step, in which the scope of negation is shifted from outside a belief to alternatives inside a belief. However, whether or not listeners arrive at exhaustive interpretations through this type of Gricean reasoning is disputed (Sauerland 2004, Spector 2007, Fox 2014, Chemla & Singh 2014).

An alternative model, the grammatical account (Chierchia 2013, Fox & Katzir 2011), holds that exhaustive interpretations of utterances do not need to depend on Gricean reasoning. Instead, they can be the product of syntactic and semantic operations (discussed more in depth in Section 1.2). In this theory, speaker information is integrated *prior* to the derivation of the SI instead of during the reasoning steps shown above. This process would be superfluous in the grammatical model because, with an informed speaker, the exhaustive interpretation (4) could be computed without it. In a case like (2), the contextual evidence would block an exhaustive interpretation such as (4). However, like in the Gricean model, deriving an ignorance implicature (5) would also require additional reasoning steps.

We conducted two mouse-tracking experiments to test differences between the two models of SIs found in Chemla & Singh 2014 (see Figure 1). We argue that intonation provides a novel test of the reasoning steps in these models because intonation signals both structural as well as epistemic information. Depending on which model, this information is integrated at different points during the derivation of SIs. Before providing more specifics, we discuss prior processing work on SIs and how intonation provides a novel test of the models.

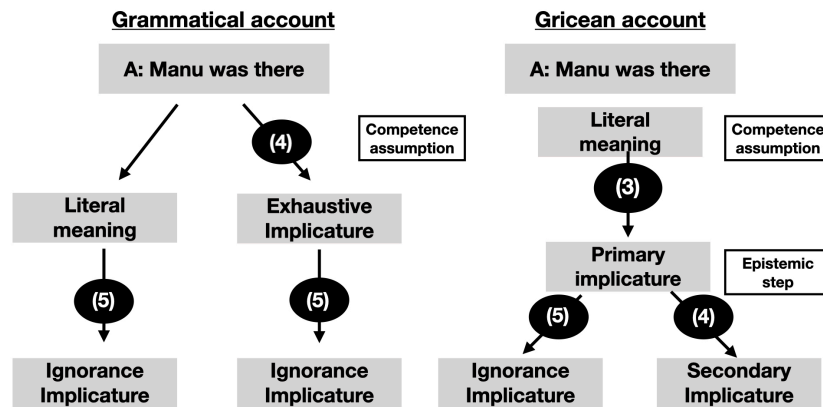


Figure 1 Interpretative steps of Gricean and grammatical accounts of SI from Chemla & Singh (2014). The numbers refer to the examples in the text.

1.1 Processing data and SIs

Recent debates on SIs have focused on whether they are more akin to conventional as opposed to conversational implicature (Levinson 2000). Several studies have found that the processing of SIs such as *some, but not all* is often delayed compared to the logical interpretation *some, possibly all*. This supports the notion that SIs involve Gricean reasoning and are cases of conversational implicatures (Bott & Noveck 2004, Breheny, Katsos & Williams 2006, Huang & Snedeker 2009, Tomlinson, Bailey & Bott 2013). Having said that, the notion of a processing cost for SIs has been disputed (Grodner et al. 2010, Degen & Tanenhaus 2015) and alternative explanations for the source of this cost have been proposed (van Tiel, Pankratz & Sun 2019).

These and other processing studies have resulted in a better understanding of how the interpretation of SIs interacts with various contextual and cognitive factors, allowing for better predictive modeling of SIs (Benz & van Rooij 2007, Franke 2009, Frank & Goodman 2012). However, this has not necessarily resulted in new theories of SIs that explicitly try to account for these data and the addition of constraints or new components to existing theories.

One reason for this might be that formal accounts of SIs are not processing theories by nature (Geurts & Rubio-Fernández 2015). However, as mentioned above, processing data have nonetheless been central in deter-

Does intonation automatically strengthen scalar implicatures?

mining the plausibility of the formal mechanisms underlying SIs. They are what Chemla & Singh (2014) have called “auxiliary assumptions”. These assumptions are a set of constraints on formal and computational models of SIs. Below we develop additional assumptions for how intonation factors into these models. They are based on prior work on intonation and implicature processing (Gotzner, Wartenburger & Spalek 2016, Tomlinson, Gotzner & Bott 2017) and we argue that these assumptions provide a reasonable test of the computational mechanisms in models of SIs.

1.2 Intonation and models of SIs

Intonation has a reputation of being a “half-tamed savage” (Bolinger 1986) because an intonational form, i.e., tone-tune combination (Pierrehumbert & Hirschberg 1990), can have multiple meanings (“one-to-many”). This presents a challenge for implicatures because intonation can carry both information about the speaker as well as grammatical aspects of the utterance. In the case of the latter, intonation often triggers linguistic focus (*F*) on a constituent, introducing a set of alternatives (Rooth 1992).

- (6) A: Were Manu and Moni at the party?
B: [MANu]_F was there → {*x* | *x* ∈ *E*, at the party}.

The constituent *Manu* would evoke a set of alternatives, namely a domain of individuals attending the party.² While theories of focus note that speakers often mark focused constituents, this is not obligatory. Regardless, intonational contours such as the H* and L+H* are strong cues for listeners to detect focus (Watson, Tanenhaus & Gunlogson 2008, Gotzner, Wartenburger & Spalek 2016, Gotzner 2019). Consider the following example:

- (7) A: Were Manu and Moni at the party?
B: [MANU]_{L+H*L-L%} was there.

Aside from reinforcing a set of “at the party” alternatives, the L+H* could help further restrict the domain of these individuals, so that the domain

² Büring (2007) argues linguistic focus does not invoke alternatives or mark contrast per se. Rather it is only supposed to highlight the relationship between a constituent receiving focus and the prior discourse context. Under this view, a specific pitch accent would only mark a constituent of its informational structural status such as “givenness”. While we do not doubt that information structural relationships play a role in SIs, we are unaware of any literature making explicit connections to SIs.

does not include *Moni*. According to the grammatical model, once an alternative stands in focus, it feeds directly into the computation of implicatures (Fox & Katzir 2011) and, unless taken up by some other operator, automatically undergoes *silent* or *covert exhaustification* (Chierchia 2013). In our case, the covert *only* operator would negate alternatives *not* in the set of focus alternatives, therefore immediately resulting in an exhaustive interpretation (e.g., *Moni*, \neg *Moni*).

Although alternative semantics and focus are powerful predictive tools for SIs, we claim that viewing intonation's role solely in this way is too narrow. In addition to structural information, intonation also provides relevant information for SIs such as speaker belief and commitments (Gunlogson 2001, Prieto 2015). In (7), the L+H* intonation signals speaker certainty (Gravano et al. 2008), further strengthening the competence assumption. Another example can be seen in (8).

- (8) A: Were Manu and Moni at the party?
 B: [MANU_{L*+HL-H%}] was there.

Here, the pattern found is known as a *rise-fall-rise* contour (Hirschberg & Ward 1992) and is traditionally analyzed as communicating degrees of speaker commitment to scalar alternatives. Unlike (7), this pitch contour asserts a *lack* of certainty on the part of the speaker about whether the unmentioned guest *Moni* attended the party. Consider this example from (Ward & Hirschberg 1988):

- (9) A: How do I get from the airport to the city center?
 B: You can take the TRAM (L*+H L-H%)

In (9), the combination of the pitch accent (L*+H) — evoking uncertainty (Ward & Hirschberg 1985) — and boundary tone (L-H%) — leaving the proposition “open” — signals that the speaker reduces *commitment* to the stronger alternative “only the tram”, resulting in an ignorance implicature. Interestingly, the rise-fall-rise can be both anchored to a specific word (“tram”) or spread out over multiple words. We mention this because listeners might first interpret the pitch accent L+H* to signal certainty or contrast, but then need to integrate that information with conflicting information from the final rising boundary tone, i.e., uncertainty. In the case of SIs, the L*+H could focus scalar alternatives (see Göbel 2019 for a discussion of L*+H and exhaustification of alternatives), though this would be canceled by the rising

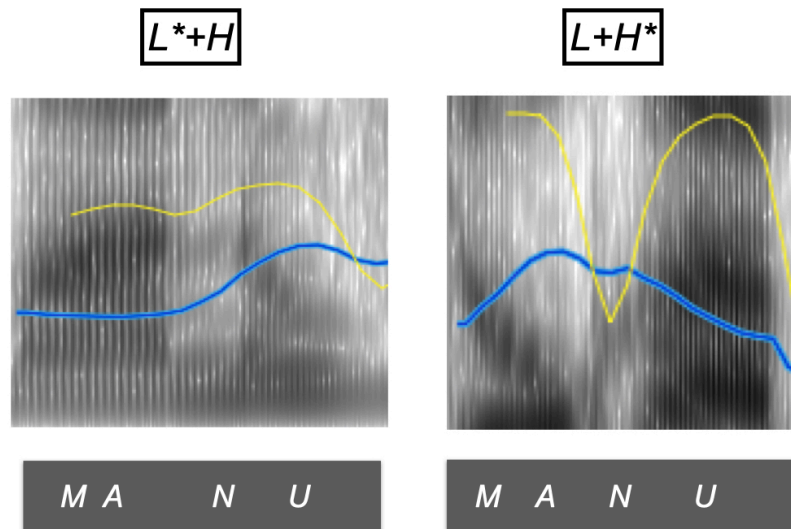


Figure 2 Pitch tracks for the L*+H and L+H* in German.

boundary tone. We see this as an issue for the grammatical theory, but not for the Gricean theory because information about speaker competence could be updated *incrementally* at multiple steps in the derivation.

From this discussion on intonation, we can derive the following predictions for the L*+H and L+H* accents (displayed in Figure 2). Under the grammatical account, exhaustive interpretations should be derived automatically if 1) the constituent stands under focus and 2) the competence assumption has been satisfied. Accordingly, if the pitch accent L+H* signals both focus as well as speaker competence, we would expect exhaustive interpretations (derived via an exhaustive-biased intonation contour) to be interpreted more immediately than ignorance implicatures (derived via an ignorance-biased intonation contour). This is because only ignorance implicatures should require additional mechanisms in the grammatical account.³ Below, we dis-

³ We thank an anonymous reviewer for pointing out that according to one version of the grammatical account, intonation could disambiguate the scope of covert exhaustification relative to a matrix “K-operator”, i.e., $\text{exh}[\text{K}[\text{Manu came}]] = \text{the speaker isn't sure that both Manu and Moni came vs. } \text{K}[\text{exh}[\text{Manu came}]] = \text{the speaker is sure that both Manu and Moni did not come}$. Because of space, we do not provide a detailed analysis of how pitch accents could serve to disambiguate these parses. We mention briefly, though, that additional operations would be needed to derive both the exhaustive (\neg Moni) and ignorance (?Moni) readings.

cuss how our experiments operationalize this prediction, namely through the L+H* and L*+H accents (Experiment 1) and by additionally introducing a competent speaker through context (Experiment 2).

2 Our investigation

Our investigation uses intonation to further test the Gricean and grammatical approaches to SIs. Most prior work has focused on lexical triggers such as *some*, *or* as cues or triggers for SIs because they explicitly assert weaker propositions and evoke stronger, even lexically entailed, alternatives. Intonation has been shown to also engage mechanisms underlying SI derivation, both for lexical triggers and bare nouns (Gotzner, Wartenburger & Spalek 2016, Gotzner 2019) and specific contours such as the L+H* can lead to quicker derivation of exhaustive interpretations (Gotzner, Wartenburger & Spalek 2016, Tomlinson, Gotzner & Bott 2017). According to the grammatical view, the immediate derivation of the exhaustive interpretation would be explained by both 1) the L+H* focusing alternatives and activating a covert operator (EXH) and/or 2) reinforcing the competence assumption prior to derivation (Gravano et al. 2008), also allowing for covert exhaustification. In the Gricean model, there should be no processing advantage for exhaustive interpretations over ignorance implicatures because the speaker information needed to *strengthen* SIs or *weaken* SIs is integrated after deriving the initial reasoning step of deriving a weak implicature. In two experiments, we test whether the L+H* results in more direct derivations of exhaustive interpretations (Experiment 1) and whether this is driven by the competence assumption (Experiment 2).

Both experiments used *mouse-tracking* (Spivey, Grosjean & Knoblich 2005, Spivey 2007, Freeman & Ambady 2010) because it provides many advantages over traditional measures like reaction times for studying higher-level cognitive processes. First, mouse-tracking effectively unpacks a button press by linking how mouse-movements can unfold over a response. Instead of a response being “fast” or “slow”, the pattern of movements found in responses can reduce the hypothesis space as to *why* a response might be effortful or

While this version of the grammatical account and the Gricean account would make similar predictions for Experiment 1, Experiment 2 explicitly manipulates speaker knowledge, effectively disambiguating the parse, and as such this version of the grammatical account would still predict higher derivation rates and earlier derivation of exhaustive interpretations with a strong competence assumption, and vice versa for ignorance implicatures

Does intonation automatically strengthen scalar implicatures?

delayed because of the *type* of movement (Spivey, Grosjean & Knoblich 2005, Wulff et al. 2019). In the case of SIs, Tomlinson, Bailey & Bott (2013) found that participants initially pushed the mouse towards weaker alternatives prior to selecting stronger alternatives, suggesting that the processing costs sometimes found with SIs result from (at least) two interpretative steps. Mouse-tracking has also provided better linking hypotheses for interpretative steps in other semantic and pragmatic phenomena such as negation (Dale & Duran 2011, Maldonado, Dunbar & Chemla 2019).

An additional advantage of mouse-tracking is the emphasis on the spatial components of a response as opposed to raw time-course. This means that, for an experimental setup involving a comparison between a ‘target’ and a ‘competitor’ option, it is possible to investigate whether delays in responses are the product of indecision or of competing (parallel) activation of responses. In our specific case, the Gricean model predicts that both exhaustive interpretations and ignorance implicatures are equally likely to occur after the first reasoning step (primary implicature). Therefore, when a pitch accent biases towards one or the other interpretation ($L+H^*$ = exhaustive bias, L^*+H = ignorance bias), the two should be derived in a similar way. The grammatical theory, on the other hand, predicts no delay for the exhaustive interpretations because $L+H^*$ should help listeners recognize focus and assess competence. Figures 3 and 4 schematically depict these predictions as well as our paradigm using one of our critical items as an example.

3 Experiment 1

3.1 Procedure

Participants first read a question (“Were Manu and Moni at the party”), then heard the answer to this question (“Manu was there”) and subsequently selected the interpretation that best matched what the speaker was trying to communicate. Participants saw five practice trials and then continued to the main experiment.

The experiment was programmed and run using the *Mousetracker* suite (Freeman & Ambady 2010). To start a trial, a participant had to click on the START button. A written question then appeared on the screen. After reading the question the participant pressed ENTER. Participants then heard an answer to the question (see Table 1) and could move their mouse to select one of the response options. Participants were allowed to start moving their

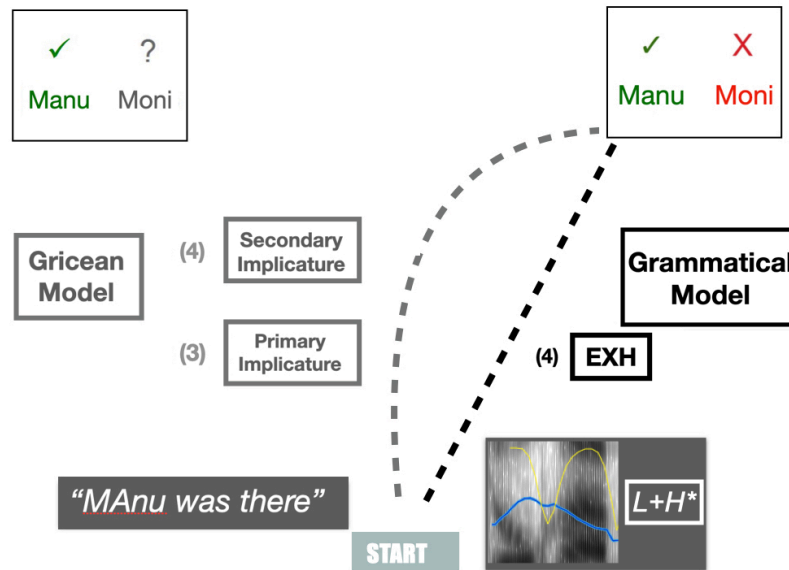


Figure 3 Example of experimental trial with predictions for the L+H* accent. Numbers refer to examples in the text.

mouse approximately 500 milliseconds after the audio stimulus started playing. This roughly amounted to the onset of the verb, *war* 'was' for the intonational conditions. A warning message appeared if participants took longer than 800ms to initiate a mouse-movement. A warning message also appeared if they took more than 2500ms to respond. Such trials were excluded from the analysis (less than 2% of all trials). Participants took anywhere from 25 to 40 minutes to complete the experiment.

3.2 Design

The Experiment had a 3X2 factorial design with the factors CUE TYPE (lexical early, lexical late and intonation) and BIAS (exhaustive vs. ignorance).

Two counterbalanced lists were made to ensure that participants saw only one version of an item for its BIAS: either its intonational pattern (L+H* or L*+H) or lexical marker (exhaustive: *nur* 'only' or *allein* 'alone' or ignorance *zumindest* 'at least' or *auf jeden Fall* 'definitely'). Cue type was distributed over different items as the goal of our study was not to test between lexical and intonational triggers. Rather, lexical conditions served as a con-

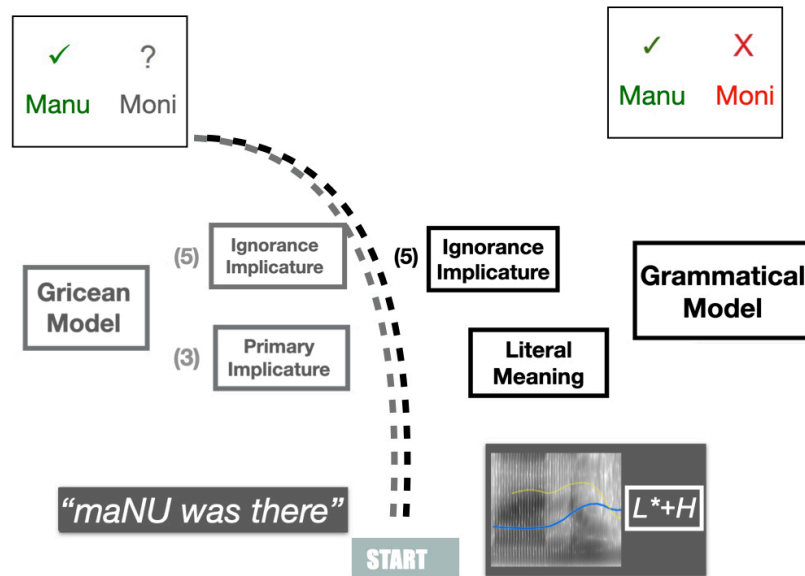


Figure 4 Example of experimental trial with predictions for the L*+H accent. Numbers refer to examples in the text.

control to the disambiguation effects of the pitch accents. Because pitch accents were distributed over the proper noun and could not disambiguate a referent until the end of the word, the *early* or *late* lexical conditions provided an additional control for the position of the trigger.

3.3 Materials

All experimental items had two response options: the exhaustive interpretation (Manu√, Moni X) and the ignorance interpretation (Manu√, Moni ?). The stimuli were taken from 6 different speakers (4 female, 2 male), and were based on the spontaneously produced utterances from Tomlinson & R. Ronderos 2014. All conditions for an example critical item are shown in Table 1.

Filler trials (40 items) had different types and combinations for response options, e.g., *Marie and Tom were both at the party* (A √, B√), *Neither A, nor B were at the party* (A X, B X), *I don't know if A or B was there* (A ?, B ?) as well as for dialogues and audio, e.g., "Was A at the party" or "Were A and B at the party" were combined with "A was there", "B was there", "A and B were there", "Neither were there".

CUE TYPE (N)	BIAS	TRIGGER	EXAMPLE
Intonation (5)	Exhaustive	L+H*	[$MANU_{L+H*L-L\%}$] <i>war da</i> 'Manu was there'
Intonation (5)	Ignorance	L*+H	[$MANU_{L*+HL-L\%}$] <i>war da</i> 'Manu was there'
Lexical early (5)	Exhaustive	<i>Nur</i> 'only'	<i>Nur Manu war da</i> 'only Manu was there'
Lexical early (5)	Ignorance	<i>Zumindest</i> 'at least'	<i>Zumindest Manu war da</i> 'at least Manu was there'
Lexical late (5)	Exhaustive	<i>Allein</i> 'alone'	<i>Manu war allein da</i> 'Manu was there alone'
Lexical late (5)	Ignorance	<i>Auf jeden Fall</i> 'definitely'	<i>Manu war auf jeden Fall da</i> 'Manu war definitely there'

Table 1 Stimuli and conditions for Experiments 1 and 2

Because our experiments were conducted in German, we examined slightly different intonational contours compared to those that have been investigated in English. Whereas the form and function of L+H* accent in German is quite comparable to that of English, the rise-fall-rise contour does not exist in German. Instead, we investigated the L*+H contour, which has been shown to signal uncertainty (Féry 1993). This contour is also minimally contrastive with L+H*. This also serves as a better experimental control because both pitch accents in German are anchored to the proper noun, whereas as mentioned in Section 1.2, the rise-fall-rise can be distributed over multiple constituents and positions over the utterance, leading to differences in when relevant information is processed.

Pitch accents were acoustically manipulated using Praat's resynthesis (LPC) functions (Boersma 2001). We attempted to maintain the tonal structure of the contours, but some manipulations were not felicitous due to faulty pitch tracks or a noticeable artificial-sounding voice. To remedy this, a German native speaker re-recorded three items as well as two items for the lexical sets and these replaced two out of the 5 items in each set. All items were scaled for intensity (67 dBs) and had similar duration over the entire utterance. These materials are available on our OSF repository <https://osf.io/85qav/>.

Does intonation automatically strengthen scalar implicatures?

3.4 Predictions

Experiment 1 tests whether pitch accents can disambiguate between exhaustive interpretations (Manu \checkmark , Moni X) and ignorance implicatures (Manu \checkmark , Moni ?), and if so, whether there are systematic differences in how this disambiguation unfolds spatially and chronologically. Based on findings from [Gotzner, Wartenburger & Spalek 2016](#), we expect exhaustive interpretations triggered by L+H* to be delayed relative to the focus particle *nur* ‘only’. According to the grammatical account, ignorance implicatures should be delayed relative to exhaustive interpretations, therefore ignorance implicatures triggered by L*+H should in turn be delayed relative to exhaustive interpretations triggered by L+H*. On the other hand, processing delays for SIs are seen as more compatible with Gricean accounts ([Bott & Noveck 2004](#), [Breheny, Katsos & Williams 2006](#), [Tomlinson, Bailey & Bott 2013](#)). As seen in [Figure 1](#), the Gricean accounts predicts that exhaustive interpretations triggered by L+H* should have a similar processing profile as ignorance implicatures triggered by L*+H.

3.5 Participants

Thirty-four native speakers of German (21 female, 13 male, ages 18-37) participated in the experiment in exchange for 7 Euros in compensation. They all reported normal or corrected-to-normal vision.

3.6 Results

3.6.1 Forced choice data

Target items were analyzed by CUE TYPE (lexical early, lexical late, and intonation) and by BIAS (exhaustive vs. ignorance). [Figure 5](#) shows the derivation rates for Experiment 1 (\checkmark ? for ignorance and \checkmark , X for exhaustive). A mixed effects binomial logistic regression model was fitted to the data using the `lme4` package ([Bates et al. 2015](#)). The model included BIAS and CUE TYPE as fixed effects (using a sum contrast coding scheme) with random intercepts and slopes by subject for BIAS and CUE TYPE and random intercepts by item. We chose this model based on the recommendations of [Barr et al. \(2013\)](#) for selecting the maximal random effects structure granted by the experimental design (the maximal model). However, since the maximal model did not converge, we removed the random slope by subjects for the interaction between

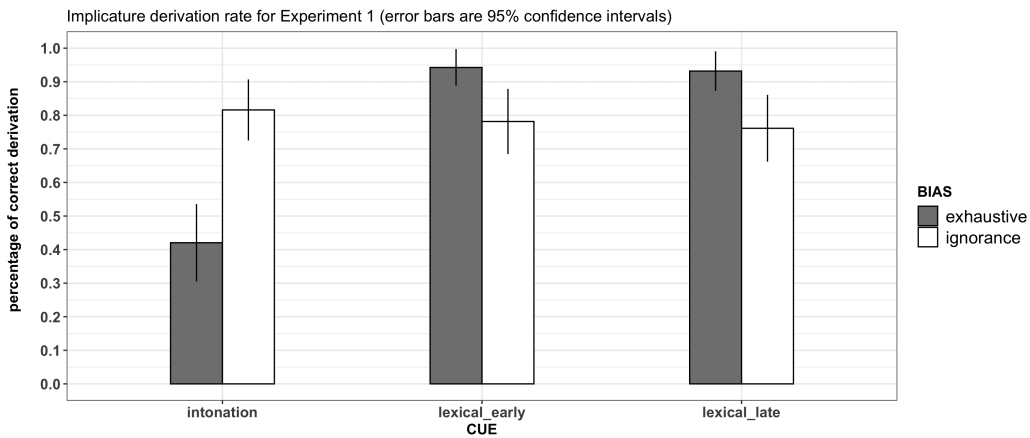


Figure 5 Implicature derivation rates for Experiment 1

BIAS and CUE TYPE. P-values were computed using a Laplace approximation. All data and code are available on the OSF repository (<https://osf.io/85qav/>).

Overall, listeners had higher derivation rates in lexical conditions asserting exhaustivity (over 90% for *only* and *alone*), whereas lexical conditions asserting ignorance (*at least* and *definitely*) had lower rates, 78% and 76% respectively, though this difference was not statistically significant. Derivation rates for the intonation conditions were substantially lower for exhaustive cases (42%) than for ignorance ones (81%) ($p < 0.05$). This cross-over interaction between and CUE TYPE can be seen in Figure 5.

FIXED EFFECTS:	ESTIMATE	STD. ERROR	Z VALUE	P-VALUE
(Intercept)	2.181	0.317	6.882	
CUE TYPE	0.561	0.385	1.457	0.145
BIAS	0.201	0.343	0.586	0.558
CUE TYPE:BIAS	0.991	0.428	2.313	<0.03*

Table 2 Experiment 1, model for derivation data

3.6.2 Mouse-tracking data pre-processing

Participants' raw data files were pre-processed using the *Mousetrap package* for R (Kieslich & Henninger 2017). The x- and y- coordinates for each trial

Does intonation automatically strengthen scalar implicatures?

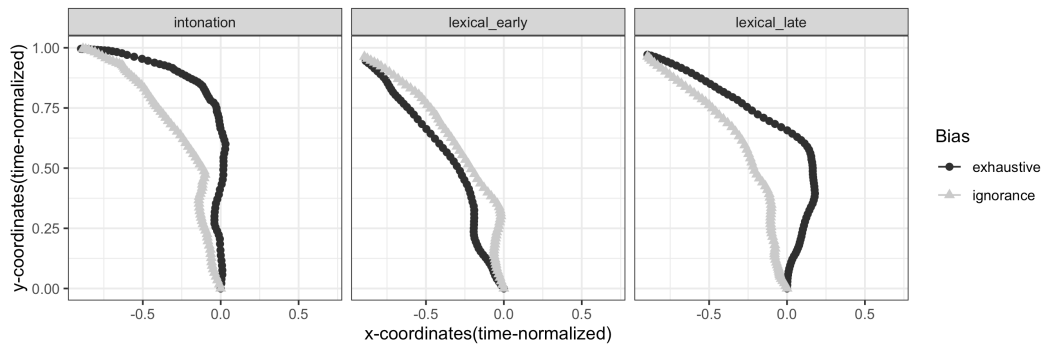


Figure 6 Average mouse trajectories across all target items in Experiment 1 separated by CUE.

were time-normalized over 101 equal time steps to analyze their overall geometrical or spatial properties. We used *area-under the curve* values (AUCs) as our dependent variable because it provides the total *overall* area over the entirety of a response, and not just at one given time-point in the response, as, for example, *maximal deviation* does.

3.6.3 Mouse-tracking data analyses

To make sure that the averaged results did not mask underlying differences in strategies between participants, the distributions of area-under-the curve values (AUCs) were analyzed for bimodality by condition using Hartigan's dip statistic (See [Freeman & Dale 2013](#) for a description). All conditions had bimodality coefficients of less than 0.45, p 's > 0.3 ., suggesting that the averaged mouse trajectories were not masking underlying distinct strategies or bimodal groups of patterns. For the analysis, we conducted one linear mixed-effects model using CUE TYPE (lexical early, intonation, and lexical late) and BIAS (exhaustive vs. ignorance). The same model was fitted again collapsing over position (lexical early vs. lexical late) to compare effects of intonation vs. lexical markings. These variables were re-coded into factors and sum-contrast coded. The model included random intercepts and slopes by subject for both main effects, as well as random intercepts by items.

Figure 6 shows the average mouse trajectories across target conditions. Listeners had more direct responses to correct targets in the lexical early condition than in the intonation condition, $t = 3.52$, $p < 0.01$. There was no

FIXED EFFECTS:	ESTIMATE	STD. ERROR	Z VALUE	P-VALUE
(Intercept)	0.299	0.045	6.591	
CUE TYPE	0.601	0.181	3.352	<0.01**
BIAS	0.087	0.061	1.403	0.171
CUE TYPE:BIAS	0.695	0.241	2.82	<0.01**

Table 3 Experiment 1, model for mouse-tracking data - AUC values

significant difference between early and lexical late conditions, $t = 1.7$, $p < 0.09$. While there was no overall main effect of BIAS on AUC values, $t = 1.4$, $p = 0.17$, there was a significant interaction between BIAS and CUE TYPE for the intonation conditions, $t = 2.82$, $p < 0.01$, albeit not for lexical late conditions, $t = 1.2$, $p = 0.23$. Details of the model can be found in Table 3.

4 Discussion

Experiment 1 investigated how pitch accents contribute to the derivation of SIs. According to the grammatical approach, a pitch accent such as L+H* should trigger a semantic operation (EXH) and result in an early derivation of a SI for the exhaustive interpretation. Surprisingly, we found quite the opposite: utterances with the L+H* had both lower implicature derivation rates and displayed longer competition with the ignorance implicature response option. The latter finding suggests that the exhaustive interpretation are immediately available.

However, this finding only partially speaks to our main research question because the speaker's knowledge state was not explicitly specified, rather only inferred from the pitch accent. Participants could have either assumed a competent speaker or equally assumed that the speaker did not have sufficient information to derive an exhaustive implicature. The latter would explain the overall preference for the ignorance response option across Experiment 1. The goal of Experiment 2 was then to address this possibility in addition to providing an explicit manipulation of speaker knowledge.

5 Experiment 2

Experiment 2 was a further test of the role of intonation in the derivation of SIs with the addition of explicitly manipulating the competence assumption. Specifically, it tests whether the contextual licensing of speaker competence

Does intonation automatically strengthen scalar implicatures?

results in exhaustive interpretations being derived prior to ignorance implicatures.

5.1 Design

The design was identical to Experiment 1 with two exceptions. The first was the addition of the factor SPEAKER (*competent vs. incompetent*). In the *competent speaker* context, listeners were told via a prompt that the speaker had stayed for the entirety of the party in question and had complete knowledge of all of the guests in attendance. In the *incompetent speaker* context, listeners read a different prompt: the speaker had only stopped by the party for a couple of minutes and had very limited knowledge regarding the attendees of the party in question. The second difference between Experiments 1 and 2 was that early and lexical late conditions were collapsed into a single condition for statistical analysis. This resulted in a 2x2x2 design with the factors CUE TYPE (lexical vs. intonational), BIAS (ignorance vs exhaustive) and SPEAKER (competent vs. incompetent).

5.2 Predictions

The predictions for Experiment 2 are similar to those in Experiment 1. The speaker manipulation should either strengthen (*competent speaker*) or weaken (*incompetent speaker*) the competence assumption. As per the grammatical account, the competent speaker manipulation should allow the L+H* to activate EXH, resulting in more direct activation for the exhaustive interpretation with L+H* in the competent speaker context than in the incompetent speaker context. As with Experiment 1, exhaustive interpretations in the grammatical account should always be derived prior to ignorance implicatures, and the speaker contexts should provide an additional test of this. The Gricean theory, while providing less clear predictions, should produce a similar pattern for exhaustive interpretations with L+H* across the speaker contexts, but would also predict that L*+H would result in less delays in the incompetent speaker context than the competent speaker context. Again, here we assume that this type of information is integrated after deriving a weak implicature (3). In other words, the speaker context manipulation should result in exhaustive interpretations being the least delayed in competent speaker context, whereas ignorance implicatures should be the least delayed in the incompetent speaker context.

5.3 Participants

Seventy-nine native speakers of German (48 female, 31 male, ages 18-39) participated in the experiment in exchange for 8 Euros in compensation. They all reported normal or corrected-to-normal vision.

5.4 Materials & stimuli

The same materials were used from Experiment 1 with the addition of a context prompt. These contexts included an avatar of the speaker, which was created using the online game ‘Simpsonsmaker’, which is no longer available. The visual prompts used in all experiments have been made available on our OSF pre-registration. The avatars were accompanied with a two-sentence description of the speaker’s attendance of a party (target items) or an unrelated anecdote of something that happened at a party (filler items).

5.5 Results

5.5.1 Forced choice data

Figure 7 shows the results for the forced-choice data of Experiment 2 and the statistical models are shown in Table 4. The statistical analysis was the same as in Experiment 1. In the final model, we included random intercepts for both subjects and items, random slope terms for the fixed effects but not for their interactions by subjects, and a random slope for context by items. This was the converging model closest to the maximal model granted by our design.

Overall, listeners were significantly more likely to derive implicatures in the lexical condition than in the intonational conditions ($p < 0.001$), but there were no main effects of SPEAKER or BIAS. The interaction between BIAS and CUE TYPE found in Experiment 1 was replicated in Experiment 2. This further supports the claim that people are more likely to choose ignorance over exhaustive interpretations when the utterances are marked by intonation, whereas the pattern is reversed when the same information is lexically marked. Critically, the interaction was modulated by SPEAKER: whereas the difference in derivation rates for exhaustive vs. ignorance implicatures was large for incompetent speaker contexts (identical to the pattern found in Experiment 1), this difference was smaller in the competent speaker context condition ($p < 0.001$).

Does intonation automatically strengthen scalar implicatures?

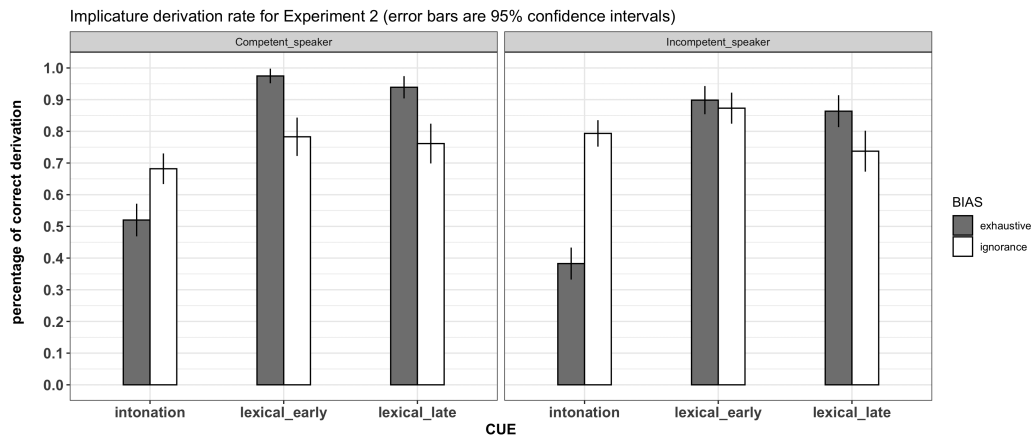


Figure 7 Implicature derivation rates for Experiment 2

FIXED EFFECTS:	ESTIMATE	STD. ERROR	Z VALUE	P-VALUE
(Intercept)	1.732	0.169	10.259	
Bias	-0.027	0.215	-0.125	0.901
Speaker	0.048	0.075	0.635	0.525
Cue Type	-1.144	0.166	-6.891	<0.001***
Bias:Speaker	0.429	0.071	6.023	<0.001***
Bias:Cue Type	-0.829	0.158	-5.253	<0.001***
Speaker:Cue Type	-0.125	0.069	-1.814	0.07
Bias:Speaker:Cue Type	-0.031	0.067	-0.46	0.646

Table 4 Experiment 2, model for derivation data

5.5.2 Mouse-tracking data

The mouse-tracking data were pre-processed and analyzed in the same way as in Experiment 1. Overall, listeners had more direct responses to correct targets for lexical conditions relative to intonation conditions, $t = 5.33$, $p < 0.001$. However, there were no main effects for either BIAS nor SPEAKER across all conditions, $t = 0.73$, $p = 0.94$ & $t = 0.62$, $p = 0.53$. Neither the interaction of BIAS and Cue type, nor the interaction of BIAS and SPEAKER, t 's < 0.8 , p 's > 0.9 , nor interaction of SPEAKER and CUE TYPE were significant, $t = 1.86$, $p = 0.83$. In contrast, the three-way interaction between CUE TYPE, BIAS, and SPEAKER was statistically significant, $t = 2.12$, $p < 0.05$.

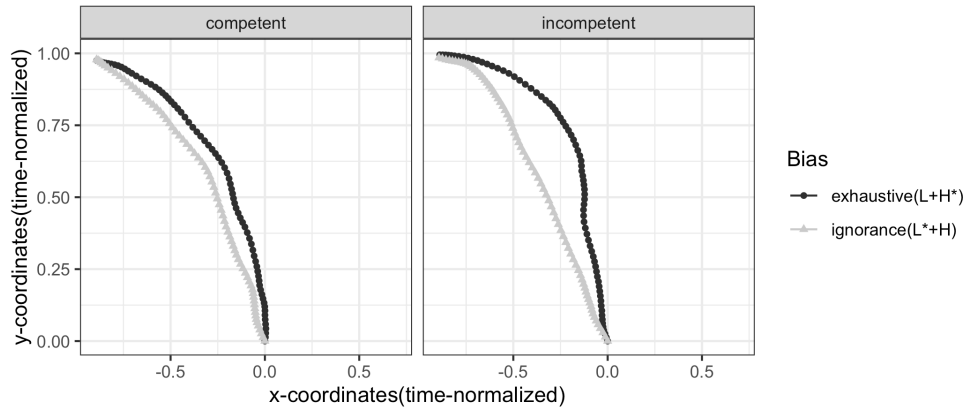


Figure 8 Average mouse trajectories for intonation conditions by speaker context in Experiment 2

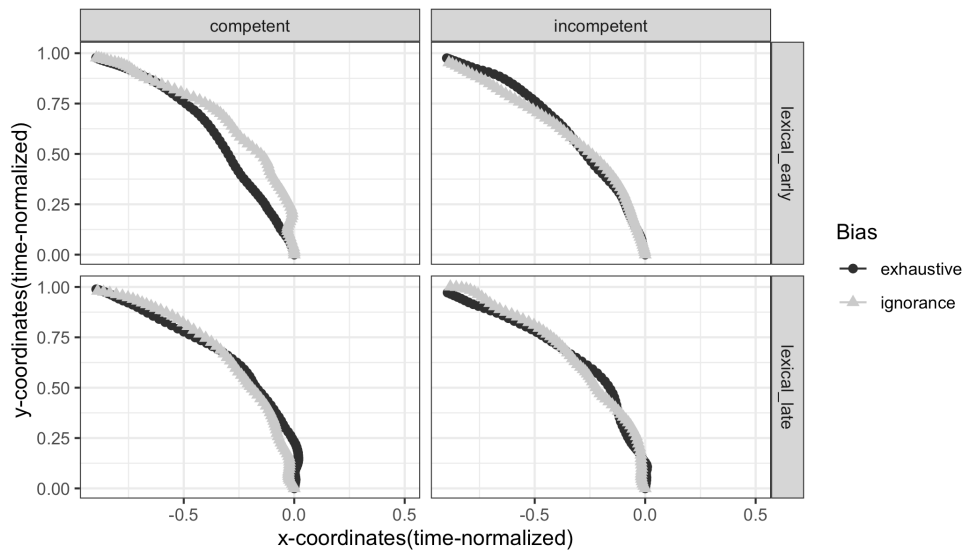


Figure 9 Average mouse trajectories for lexical conditions by speaker context and position(cue) in Experiment 2

FIXED EFFECTS:	ESTIMATE	STD. ERROR	Z VALUE	P-VALUE
(Intercept)	0.337	0.008	3.31	
Bias	0.065	0.002	0.73	0.94
Type	0.955	0.053	5.33	<0.001***
Speaker	0.0225	0.008	0.62	0.53
Speaker: Type	0.173	0.0036	1.88	0.08
Bias:Type	0.0036	0.004	0.08	0.94
Bias:Speaker	0.0023	0.009	0.11	0.92
Bias:Speaker:Type	0.206	0.036	2.12	<0.04*

Table 5 Experiment 2, model for mouse-tracking data - AUC values

We also conducted a post-hoc subset analysis on the intonation conditions to examine the source of the three-way interaction. Here, there was a main effect of SPEAKER on AUC values, $t = 2.45$, $p < 0.05$, but not a significant main effect of BIAS, $t = 0.92$, $p = 0.35$. Critically, there was a significant interaction of SPEAKER and BIAS on AUC values, $t = 2.11$, $p < 0.05$. This suggests that exhaustive interpretations were delayed in incompetent contexts relative to competent contexts, as shown in Figure 8.

5.6 Discussion

Experiment 2 replicated the results of Experiment 1, showing a bias towards ignorance interpretations over exhaustive interpretations even in contexts that reinforce the competence assumption. The addition of speaker competence contexts helped rule out potential confounds with Experiment 1, i.e., the low derivation rate of exhaustive interpretations. While visual inspection of Figure 8 seems to suggest a smaller difference between the L+H* and L*+H across the speaker contexts, this seems to be due to an increased delay of the L*+H in the competent speaker condition and not more direct responses for the L+H* in this condition.

6 General Discussion

Our two mouse-tracking studies suggest that exhaustive interpretations might not be immediately available in the initial steps of implicature computation. Instead, Experiment 1 showed that ignorance implicatures were preferred, resulting in the exhaustive interpretation being derived later in time.

Experiment 2 replicated this finding even when the speaker’s competence assumption was strengthened via context. At first glance, the results appear more compatible with a Gricean account and less so with grammatical accounts. Grammatical accounts state that strong SIs should be derived during initial stages, resulting in a bias towards strong SIs when sufficient information about the speaker’s knowledge state is provided *a priori*.

However, the data from our experiments cannot definitively adjudicate between these accounts. First, neither account is a processing theory, meaning that additional linking assumptions are needed in both cases in order to explain processing data (Chemla & Singh 2014). Second, it is also possible that the bias for ignorance responses could be brought on by an informational overlap between weak implicatures and ignorance implicatures with the ignorance response targets (Manu√, Moni ?). But even if these interpretations were not mutually exclusive in relation to the ignorance response target, this would also support the Gricean model because the grammatical model has no such thing as weak implicatures, rather only delayed ignorance implicatures. Below, we discuss some limitations of our study and our contribution to theoretical developments of SIs.

6.1 Limitations of our experiments

One surprising finding was that listeners’ derivation rates for exhaustive (strong) implicatures were low in Experiment 1. We found, however, that rates increased substantially in Experiment 2 in the competent speaker contexts. This suggests that, in the absence of explicit information, participants might have assumed an uninformed speaker. This finding is relevant for other experimental work on SIs without explicit contextual manipulation: participants might fill in the contextual gaps with their own default assumptions.

An anonymous reviewer suggested an alternative explanation for our findings. Specifically, the L+H* accent might be more ambiguous than the L*+H because the L+H* seems to encompass multiple semantic and discourse functions such as contrast and discourse “new” information (Watson, Tanenhaus & Gunlogson 2008, Büring 2003). While this could be true for Experiment 1, this cannot explain the results in Experiment 2 for the incompetent speaker context (78 vs. 65% derivation rates). Likewise, the L*+H also has multiple functions in German, e.g., as a “calling contour” (Quiroz & Żygis 2017).

Does intonation automatically strengthen scalar implicatures?

We also see other reasons why processing delays found for L+H* in our study are not due to it having multiple functions. In Experiment 2, both the L+H* as well as *nur* ‘only’ showed significant differences across contexts. In fact, our findings replicate work by [Gotzner, Wartenburger & Spalek \(2016\)](#) and [Gotzner \(2019\)](#), who have already shown that L+H* in German is delayed for exhaustive interpretations relative to *nur* ‘only’. In addition, [Gotzner, Wartenburger & Spalek \(2016\)](#) found that the L+H* helps activate alternatives *not* already activated in the existing context, but both this study as well as [Husband & Ferreira \(2016\)](#) found that the second function of the L+H* is to select contextually-appropriate alternatives. In our study, all relevant alternatives were activated via the response prompts, hence the L+H* could have only played a role in selecting amongst the alternatives. In addition, these alternatives are *contrastive* relative to the models in question. In other words, the prior research, in both English and German, does not support the alternative explanation that a processing delay would arise from ambiguity associated with the L+H*. Rather, we find it more likely that the delay is due to the availability of secondary and ignorance implicatures *after* deriving a weak implicature.

A further reason not to attribute the delay for L+H* to ambiguity comes from examining the lexical conditions. While we see a processing advantage for *nur* ‘only’ in the competent speaker context, it is not the case that we see a corresponding processing advantage for *zumindest* ‘at least’ in the incompetent context. Put differently, why would the speaker context affect lexical assertions of exhaustivity, but not lexical assertions of ignorance? One possibility might well be that the experimental setup introduced a response bias towards ignorance responses, but that seems not to be these case in Experiment 2 for the *nur* ‘only’ items. These questions certainly warrant further experimental investigation.

6.2 Limitations of current models of SIs

We interpret our findings as showing that models of scalar implicature require a *dynamic* update of speaker meaning. In particular, because intonational meaning is distributed over the entire sentence, any parse would need a way to input intonational cues to changes in speaker meaning at multiple points of the interpretative process. We argue that the epistemic step model ([Sauerland 2004](#)) is the most plausible framework for this because speaker information can enter into the parse at multiple points.

We see current models of SIs as being limited in their ability to integrate both linguistic as well as speaker information from intonational contours. We argue that models of SIs require a dynamic update mechanisms for speaker information because intonation unfolds incrementally during the parse. Only the Gricean model discussed in the introduction allows these updates at multiple points in the derivation. The issue with the grammatical account is that it requires intonational meaning to be first converted to syntactic operators. Once the parse has started, there is no mechanism to dynamically update speaker information. This doesn't mean that the grammatical theory is incapable of doing this, rather current iterations of it lack the specifics needed to test the efficacy of the model for intonation.

6.3 Towards a processing model of SIs

Our experiments show that intonation, in this case the L+H*, does not result in automatic SI derivations for exhaustive interpretations, even with a strong competence assumption. While this does not unequivocally support one account over the other, it does suggest that more interpretative steps are needed to derive exhaustive and ignorance implicatures than are postulated in the grammatical model.

How should processing data inform predictions made by formal accounts of SIs? A common move is to appeal to Marr's levels (Marr 1982). Formal and computational accounts exist on the computational level ("what"), whereas processing data belong at the algorithmic level ("how"). According to this argument, questions about processing and theory exist only at different levels and data from each level have no bearing on the other (Geurts & Rubio-Fernández 2015). Marr (1982), however, posited an additional level for vision (called the 'interactive level'), in which levels influence and constrain each other. Any complete theory, in his view, should have explanations (and not just instantiations) at multiple levels, together with a description of how the different levels constrain one another. We see no reason why this should not extend to language as well.

To conclude, formal models of SIs do entail processing assumptions even if they are not explicit about these. Our study adds to many others that show how processing data can inform and test the architectures of such models (Bott & Noveck 2004, Breheny, Katsos & Williams 2006, Huang & Snedeker 2009, Grodner et al. 2010, Degen & Tanenhaus 2015). Despite this, processing data seems to have become less central to theory building for SIs than gram-

Does intonation automatically strengthen scalar implicatures?

mational judgments and off-line derivational data. As mentioned, this might be because linking hypotheses between processing theory and data have not been adequately fleshed out. While we have attempted to provide reasonable auxiliary assumptions and fair tests of these models, the job of defining processing assumptions for formal models of SI should not fall solely on experimentalists. We therefore call on theoreticians to play their part in realizing one of the central goals of Experimental Pragmatics: to build theories that can explain and predict SIs across multiple types of data.

References

- Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255-278. <https://doi.org/10.1016/j.jml.2012.11.001>.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1-48. <https://doi.org/10.18637/jss.v067.i01>.
- Benz, Anton & Robert van Rooij. 2007. Optimal assertions, and what they implicate. A uniform game theoretic approach. *Topoi* 26(1). 63-78. <https://doi.org/10.1007/s11245-006-9007-3>.
- Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5(9/10). 341-345.
- Bolinger, Dwight. 1986. *Intonation and its parts: Melody in spoken English*. Stanford, CA: Stanford University Press. <https://doi.org/10.1515/9781503622906>.
- Bott, Lewis & Emmanuel Chemla. 2016. Shared and distinct mechanisms in deriving linguistic enrichment. *Journal of Memory and Language* 91. New Approaches to Structural Priming, 117-140. <https://doi.org/10.1016/j.jml.2016.04.004>.
- Bott, Lewis & Ira A. Noveck. 2004. Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language* 51(3). 437-457. <https://doi.org/10.1016/j.jml.2004.05.006>.
- Breheny, Richard, Heather J. Ferguson & Napoleon Katsos. 2013. Taking the epistemic step: Toward a model of on-line access to conversational implicatures. *Cognition* 126(3). 423-440. <https://doi.org/10.1016/j.cognition.2012.11.012>.

- Breheny, Richard, Napoleon Katsos & John Williams. 2006. Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition* 100(3). 434-463. <https://doi.org/10.1016/j.cognition.2005.07.003>.
- Büring, Daniel. 2003. On d-trees, beans, and B-accent. *Linguistics and Philosophy* 26(5). 511-545. <https://doi.org/10.1023/A:1025887707652>.
- Büring, Daniel. 2007. Semantics, intonation and information structure. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, chap. 14, 445-474. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199247455.013.0015>.
- Chemla, Emmanuel. 2008. An Epistemic Step for Anti-Presuppositions. *Journal of Semantics* 25(2). 141-173. <https://doi.org/10.1093/jos/ffm017>.
- Chemla, Emmanuel & Raj Singh. 2014. Remarks on the experimental turn in the study of scalar implicature, part II. *Language and Linguistics Compass* 8(9). 387-399. <https://doi.org/10.1111/lnc3.12080>.
- Chierchia, Gennaro. 2013. *Logic in grammar: Polarity, free choice, and intervention*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199697977.001.0001>.
- Dale, Rick & Nicholas D. Duran. 2011. The cognitive dynamics of negated sentence verification. *Cognitive Science* 35(5). 983-996. <https://doi.org/10.1111/j.1551-6709.2010.01164.x>.
- Degen, Judith & Michael K. Tanenhaus. 2015. Processing scalar implicature: A constraint-based approach. *Cognitive science* 39(4). 667-710. <https://doi.org/10.1111/cogs.12171>.
- Dieuleveut, Anouk, Emmanuel Chemla & Benjamin Spector. 2019. Distinctions between primary and secondary scalar implicatures. *Journal of Memory and Language* 106. 150-171. <https://doi.org/10.1016/j.jml.2019.02.008>.
- Féry, Caroline. 1993. *German intonational patterns* (Linguistische Arbeiten 285). Tübingen: Niemeyer. <https://doi.org/10.1515/9783111677606>.
- Fox, Danny. 2014. Cancelling the maxim of quantity: Another challenge for a Gricean theory of scalar implicatures. *Semantics and Pragmatics* 7(5). 1-20. <https://doi.org/10.3765/sp.7.5>.
- Fox, Danny & Roni Katzir. 2011. On the characterization of alternatives. *Natural Language Semantics* 19(1). 87-107. <https://doi.org/10.1007/s11050-010-9065-3>.

Does intonation automatically strengthen scalar implicatures?

- Frank, Michael C. & Noah D. Goodman. 2012. Predicting pragmatic reasoning in language games. *Science* 336(6084). 998–998. <https://doi.org/10.1126/science.1218633>.
- Franke, Michael. 2009. *Signal to act: Game theory in pragmatics*. ILLC Dissertation Series: Universiteit van Amsterdam Ph.D Thesis.
- Freeman, Jonathan B. & Nalini Ambady. 2010. MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods* 42(1). 226–241. <https://doi.org/10.3758/BRM.42.1.226>.
- Freeman, Jonathan B. & Rick Dale. 2013. Assessing bimodality to detect the presence of a dual cognitive process. *Behavior Research Methods* 45(1). 83–97. <https://doi.org/10.3758/s13428-012-0225-x>.
- Geurts, Bart & Paula Rubio-Fernández. 2015. Pragmatics and processing. *Ratio* 28(4). 446–469. <https://doi.org/10.1111/rati.12113>.
- Göbel, Alexander. 2019. Additives pitching in: L*+h signals ordered focus alternatives. *Semantics and Linguistic Theory (SALT)* 29. 279–299. <https://doi.org/10.3765/salt.v29i0.4612>.
- Gotzner, Nicole. 2019. The role of focus intonation in implicature computation: A comparison with *only* and *also*. *Natural Language Semantics* 27(3). 189–226. <https://doi.org/10.1007/s11050-019-09154-7>.
- Gotzner, Nicole, Isabell Wartenburger & Katharina Spalek. 2016. The impact of focus particles on the recognition and rejection of contrastive alternatives. *Language and Cognition* 8(1). 59–95. <https://doi.org/10.1017/langcog.2015.25>.
- Gravano, Agustín, Stefan Benus, Julia Hirschberg, Elisa Sneed German & Gregory Ward. 2008. The effect of contour type and epistemic modality on the assessment of speaker certainty. *Speech Prosody (SP 2008)*. 401–404.
- Grice, H. Paul. 1975. Logic and conversation. In *Speech acts* (Syntax and Semantics 3), 41–58. https://doi.org/10.1163/9789004368811_003.
- Grodner, Daniel J., Natalie M. Klein, Kathleen M. Carbary & Michael K. Tanenhaus. 2010. “some,” and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition* 116(1). 42–55. <https://doi.org/10.1016/j.cognition.2010.03.014>.
- Gunlogson, Christine A. 2001. *True to form: Rising and falling declaratives as questions in English*. Santa Cruz, CA: University of California, Santa Cruz dissertation.
- Hirschberg, Julia & Gregory Ward. 1992. The influence of pitch range, duration, amplitude and spectral features on the interpretation of the rise-

- fall-rise intonation contour in English. *Journal of Phonetics* 20(2). 241-251. [https://doi.org/10.1016/S0095-4470\(19\)30625-4](https://doi.org/10.1016/S0095-4470(19)30625-4).
- Hirschberg, Julia B. 1985. *A theory of scalar implicature*. Philadelphia, PA: University of Pennsylvania dissertation.
- Huang, Yi Ting & Jesse Snedeker. 2009. Online interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Cognitive Psychology* 58(3). 376-415. <https://doi.org/10.1016/j.cogpsych.2008.09.001>.
- Husband, E. Matthew & Fernanda Ferreira. 2016. The role of selection in the comprehension of focus alternatives. *Language, Cognition and Neuroscience* 31(2). 217-235. <https://doi.org/10.1080/23273798.2015.1083113>.
- Kieslich, Pascal J. & Felix Henninger. 2017. Mousetrap: An integrated, open-source mouse-tracking package. *Behavior Research Methods* 49(5). 1652-1667. <https://doi.org/10.3758/s13428-017-0900-z>.
- Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/5526.001.0001>.
- Maldonado, Mora, Ewan Dunbar & Emmanuel Chemla. 2019. Mouse tracking as a window into decision making. *Behavior Research Methods* 51(3). 1085-1101. <https://doi.org/10.3758/s13428-018-01194-x>.
- Marr, David. 1982. *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman & Company.
- Pierrehumbert, Janet. & Julia. Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in communications*, 271-311. Cambridge, MA: The MIT Press.
- Prieto, Pilar. 2015. Intonational meaning. *WIREs Cognitive Science* 6(4). 371-381. <https://doi.org/10.1002/wcs.1352>.
- Quiroz, Sergio I. & Marzena Żygis. 2017. The vocative chant and beyond: German calling melodies under routine and urgent contexts. *Interspeech 2017, 18th annual conference of the International Speech Communication Association*. 1208-1212. <https://doi.org/10.21437/Interspeech.2017-1635>.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75-116. <https://doi.org/10.1007/BF02342617>.
- Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27(3). 367-391. <https://doi.org/10.1023/B:LING.0000023378.71748.db>.

- Spector, Benjamin. 2007. Scalar implicatures: Exhaustivity and Gricean reasoning. In Maria Aloni, Alastair Butler & Paul Dekker (eds.), *Questions in dynamic semantics* (Current Research in the Semantics / Pragmatics Interface 17), 225-249. Elsevier. https://doi.org/10.1163/9780080470993_011.
- Spivey, Michael. 2007. *The continuity of mind*. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195170788.001.0001>.
- Spivey, Michael J., Marc Grosjean & Günther Knoblich. 2005. Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences* 102(29). 10393-10398. <https://doi.org/10.1073/pnas.0503903102>.
- Tomlinson, John M., Todd M. Bailey & Lewis Bott. 2013. Possibly all of that and then some: Scalar implicatures are understood in two steps. *Journal of Memory and Language* 69(1). 18-35. <https://doi.org/10.1016/j.jml.2013.02.003>.
- Tomlinson, John M., Nicole Gotzner & Lewis Bott. 2017. Intonation and pragmatic enrichment: How intonation constrains ad hoc scalar inferences. *Language and Speech* 60(2). 200-223. <https://doi.org/10.1177/0023830917716101>.
- Tomlinson, John Michael & Camilo R. Ronderos. 2014. The production of partial answers and ad-hoc inferences: Evidence from semi-spontaneous speech. *Formal & Experimental Pragmatics Workshop*. 51-55.
- van Rooij, Robert & Katrin Schulz. 2004. Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information* 13(4). 491-519. <https://doi.org/10.1007/s10849-004-2118-6>.
- van Tiel, Bob, Elizabeth Pankratz & Chao Sun. 2019. Scales and scalarity: Processing scalar inferences. *Journal of Memory and Language* 105. 93-107. <https://doi.org/10.1016/j.jml.2018.12.002>.
- Van Tiel, Bob, Emiel Van Miltenburg, Natalia Zevakhina & Bart Geurts. 2014. Scalar diversity. *Journal of Semantics* 33(1). 137-175. <https://doi.org/10.1093/jos/ffu017>.
- Ward, Gregory & Julia Hirschberg. 1985. Implicating uncertainty: The pragmatics of fall-rise intonation. *Language* 61(4). 747-776. <https://doi.org/10.2307/414489>.
- Ward, Gregory & Julia Hirschberg. 1988. Intonation and propositional attitude: The pragmatics of L*+ HLH%. *Eastern States Conference on Linguistics (ESCOL)*. 512-522.

- Watson, Duane G., Michael K. Tanenhaus & Christine A. Gunlogson. 2008. Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cognitive Science* 32(7). 1232-1244. <https://doi.org/10.1080/03640210802138755>.
- Wulff, Dirk U., Jonas M. B. Haslbeck, Pascal J. Kieslich, Felix Henninger & Michael Schulte-Mecklenbeck. 2019. Mouse-tracking: Detecting types in movement trajectories. In Michael Schulte-Mecklenbeck, Anton Kühberger & Joseph G. Johnson (eds.), *A Handbook of Process Tracing Methods*, 2nd edn., 131-145. New York: Routledge. <https://doi.org/10.4324/9781315160559-10>.

John Michael Tomlinson, Jr.
Leibniz Centre for
General Linguistics (ZAS)
Schützenstrasse 18
10117 Berlin
Germany
otomlins@gmail.com

Camilo R. Ronderos
University of Oslo
Institutt for filosofi, idé- og kunsthistorie
og klassiske språk
Blindernveien 31 Georg Morgenstiernes
hus
0313 Oslo
Norway
camilorr@uio.no