Top-down information is more important in noisy situations: Exploring the role of pragmatic, semantic, and syntactic information in language processing

Fabio Trecca (fabio@cc.au.dk)
School of Communication and Culture, Aarhus University, 8000 Aarhus, Denmark

Kristian Tylén (kristian@cc.au.dk)
Riccardo Fusaroli (fusaroli@cas.au.dk)
School of Communication and Culture & Interacting Minds Centre, Aarhus University, 8000 Aarhus, Denmark

Christer Johansson (christer.johansson@uib.no)
Department of Linguistics, Literary and Aesthetic Studies, University of Bergen, 5020 Bergen, Norway

Morten H. Christiansen (christiansen@cornell.edu)
Department of Psychology, Cornell University, Ithaca, NY 14853
School of Communication and Culture & Interacting Minds Centre, Aarhus University, 8000 Aarhus, Denmark

Abstract

Language processing depends on the integration of bottom-up information with top-down cues from several different sources—primarily our knowledge of the real world, of discourse contexts, and of how language works. Previous studies have shown that factors pertaining to both the sender and the receiver of the message affect the relative weighting of such information. Here, we suggest another factor that may change our processing strategies: perceptual noise. We hypothesize that listeners weight different sources of top-down information more in situations of perceptual noise than in noise-free situations. Using a sentence-picture matching experiment with four forced-choice alternatives, we show that degrading the speech input with noise compels the listeners to rely more on top-down information in processing. We discuss our results in light of previous findings in the literature, highlighting the need for a unified model of spoken language comprehension in different ecologically valid situations, including under noisy conditions.

Keywords: sentence processing; perceptual noise; pragmatic context; real-world semantics; rational inference.

Introduction

Language processing is based on the integration of bottom-up and top-down information (Marslen-Wilson, 1987; McClelland & Elman, 1986). As we process language, the incoming input is integrated with our existing knowledge—of the local discourse contexts, of the world, and of language—and creates a frame of reference for what comes next (Ferreira & Chantavarin, 2018). This integration happens rapidly (Christiansen & Chater, 2016) and entails that the available evidence must be promptly weighted against prior information, in an effort to determine the likelihood of different specific interpretations of the perceived input (e.g., Gibson, Bergen, & Piantadosi, 2013; Levy, 2008). Success in processing is therefore dependent on the availability of reliable (probabilistic) cues to correct sentence interpretation (Martin, 2016).

At least three sources of information seem to concurrently constrain this inferential process (Venhuizen, Crocker, & Brouwer, 2019). At a local level, the syntactic structure of the language input affects the interpretation of the content of a given linguistic input. An example hereof is that the meaning of syntactically complex sentences is more likely to be misconstrued than that of their less complex counterparts: for instance, listeners more often fail to identify semantic roles in passive sentences than in active sentences (Ferreira, 2003). It has also been shown that listeners tend to take the content of semantically implausible sentences at face value when their syntactic structure is relatively straightforward (e.g., prepositional datives: *The mother gave the daughter to the candle*), but prefer more semantically plausible interpretations when the syntactic structure of the sentences is more complex (e.g., the double-object dative *The mother gave the candle the daughter* is misread as *The mother gave the candle to the daughter*)—even if the semantic content of the two sentences is identical (Gibson et al., 2013).

Lexical-semantic information rooted in our 'real-world' knowledge also points toward specific interpretations of the linguistic input and can even overrule syntactic information (see e.g., MacDonald, Pearlmutter, & Seidenberg, 1994). Semantic properties of the constituents of a sentence, such as animacy, have been shown to affect the inferential process: for instance, listeners tend to interpret animate characters as agents in *who-did-what-to-whom* sentences, independently of syntax (e.g., Larsen & Johansson, 2008; Szewczyk & Schriefers, 2011). This animate-agency bias is consistent with the suggestion that our semantic knowledge may largely originate from sensorimotor representations (see e.g., *situation model* theories of sentence processing; e.g., Zwaan, 2016), which drives listeners toward interpretations of the input that fit with their knowledge of state of affairs in the real world (e.g., Fillenbaum, 1974).

Lastly, the broader discourse context in which a given linguistic input is embedded can affect (and even overrule) our interpretation of semantic and syntactic cues.

Referential/pragmatic contexts and lexical semantics seem to have an additive influence on processing, with (linguistic and extralinguistic) contextual information playing a central role in disambiguating syntactical ambiguities (e.g., the sentence *put the apple on the napkin in the box*, in which the listener can disambiguate whether *on the napkin* modifies *the apple* or *in the box* only by relying on the informativeness of, e.g., elements in the visual world; Snedeker & Trueswell, 2004; see also Spivey, Tanenhaus, Eberhard, & Sedivy, 2002). Pragmatic/contextual expectations can even override our semantic preference for animate agents, for instance through the introduction of a discourse context in which an inanimate object is presented as the agent: Nieuwland and Van Berkum (2006) showed that animacy violations (e.g., *The peanut was in love*), which normally elicit clear N400 effects in ERP experiments, do not do so when the sentences are presented in a context that justifies the violation (e.g., *A woman saw a dancing peanut who had a big smile on his face.* […] *The peanut was in love*). In these semantically implausible contexts, the more canonical sentences (e.g., *The peanut was salted*) suddenly become the violation to the pragmatic/contextual expectations.

All three information sources—pragmatic/contextual information, real-world semantics, and syntax—converge ideally to determine one unequivocal interpretation of the input (cf. Bates & MacWhinney, 1989). However, the relative weighting of each of these information sources in different processing situations seems to be affected by properties of the language input, as well as of the language users. For instance, Dąbrowska and Street (2006) showed that demographic factors such as years of formal education predicted the listeners' ability to interpret semantically implausible sentences when these were presented in passive constructions (e.g., *The soldier was protected by the boy*). Less educated listeners tended to disregard syntactic cues and focus more on semantic and pragmatic/contextual cues (e.g., interpreting the sentence as the more plausible *The soldier protected the boy*). Similar observations have been made in relation to language spoken by non-native speakers: for instance, Gibson et al. (2017) showed that English speakers were more likely to accept literal interpretations of semantically implausible sentences, if these were produced by native English speakers, than if the speakers talked with a foreign accent (thus giving foreigners the benefit of the doubt). Likewise, both children and adults have been shown to adjust their weighting of cues based on the apparent reliability of cues in the input, for instance by being more willing to accept implausible sentences from speakers who previously have produced more implausible utterances (Yurovsky, Case, & Frank, 2017; see also Gibson et al., 2013).

In this study, we suggest that factors pertaining to the communicative environment—e.g., the presence of perceptual noise—are also likely to affect the dynamic weighting of different information sources. The aim of the present study is therefore two-fold: First, we devise a novel experimental paradigm that allows us to individuate and access the relative weight given to different sources of information (pragmatic context, semantics, and syntax) in language processing. Second, we investigate how these weights are dynamically shifted relative to each other as a function of extra-linguistic conditions that can hinder speech communication—in this case, acoustic noise in the speech signal.

Language processing in the real world is prone to be affected by noise (Shannon, 1948): conversations in crowded places or phone calls with bad reception are but a few examples of how noise commonly affects language use in everyday situations (see Mattys, Davis, Bradlow, & Scott, 2012). In these situations, listeners have been shown to devote more cognitive effort to compensate for the reduced informativeness of the signal (Peelle, 2018). Here, we propose that, in order to compensate for less informative bottom-up input, listeners dynamically shift how they weight different information sources: in situations of noise, we are more likely to rely less on bottom-up information and implicitly adopt a more top-down-guided processing style. To test this hypothesis, we used a simple sentence-picture matching task to probe for comprehension. Participants listened to eight short stories; after each story, the participants were presented with four pictures in a four-alternative forced-choice (4AFC) test and instructed to select the picture that matched the central event of the story. In each 4AFC test, only one picture matched the actual language input; the three remaining pictures corresponded to different potential misinterpretations of the language input, and they were specifically designed to reveal processing biases driven by one or more of the three information sources under scrutiny. Half of the participants listened to the short stories in a baseline condition without noise; the other half was presented with the same stories under conditions of perceptual noise.

Method

Participants

167 native Norwegian-speaking (56% female; age: $M = 23.4$, $SD = 3.03$), right-handed undergraduate and graduate students from the University of Bergen (Bergen, Norway) participated in exchange for monetary compensation. Participants were pre-screened for previous or current neurological and/or psychiatric diagnoses, dyslexia, and hearing impairments. The participants were randomly assigned to two experimental conditions: Noise and No-noise ($N_{noise} = 89$, $N_{no\text{-}noise} = 78$).

Materials

Speech stimuli The language stimuli were eight aurally-presented short stories. All stories had an identical narrative structure consisting of four sentences, as in the following example (approximate translation from Norwegian):

S1: The boy walked into the pet store.
S2: His younger sister had been wanting a goldfish for a long time, and now it was time for her to get one.

S3: Everybody thought it was adorable that the boy bought a goldfish for his sister.
S4: As expected, his sister was very happy.

S1 and S2 provided the pragmatic context of the story; S3 was the target sentence and contained the central event of the story (underlined in the example), which was to be matched to the relevant image; and S4 served as a wrap-up sentence. All stories comprised three characters: an agent (e.g., the boy), an object (e.g., the goldfish), and a recipient (e.g., the sister). By switching roles between agent and object, we created different versions of each story, in which both the pragmatic context (S1+S2) and the central event of the story (S3) could be either plausible or implausible in relation to real-world semantics (e.g., S1: *the boy walked into the pet store* vs. *the goldfish walked into the pet store*; S3: *[…] the boy bought a goldfish for his sister* vs. *the goldfish bought a boy for its sister*). Additionally, we manipulated the markedness of the syntactic structure of the target sentence in S3, so that the main event was expressed either using a prepositional dative (unmarked, e.g., *the boy bought a goldfish for his sister*) or a double object construction (marked, e.g., *the boy bought his sister a goldfish*). Together, these 2×2×2 manipulations (pragmatic context semantics × central event semantics × syntactic markedness) resulted in eight possible versions of each story, as shown in Table 1. Participants were tested on all eight story structures. Each story structure-type was randomly assigned to a specific story-token for each participant, so that participants only heard one version of each of the eight stories (e.g., Participant 1 heard Story 1 version A, Story 2 version B, etc.; Participant 2 heard Story 1 version B, Story 2 version C, etc.). The eight stories were interspersed with eight stories from another experiment (with an identical procedure), which served as filler trials.

Table 1: The eight possible narrative structures of Story 1

	S1+S2: Plausible	S1+S2: Implausible	
S3: Unmarked syntax	Story 1a Story 1c	Story 1b Story 1d	S3: Plausible S3: Implausible
S3: Marked syntax	Story 1e Story 1g	Story 1f Story 1h	S3: Plausible S3: Implausible

The 64 sound files (8 stories × 8 story structures) were recorded in a soundproof booth by a male native speaker of Norwegian from the Stavanger area, using an Audio-Technica AT2020 Cardioid Condenser USB microphone and Audacity version 2.2.2 for Mac. For the participants in the Noise group, Brownian noise with a signal-to-noise ratio of −19 was added to the sound files using the MixSpeechNoise function from the *praat-semiauto-master* package (https://github.com/drammock/praat-semiauto) in Praat version 6.0.31 (Boersma, 2001).

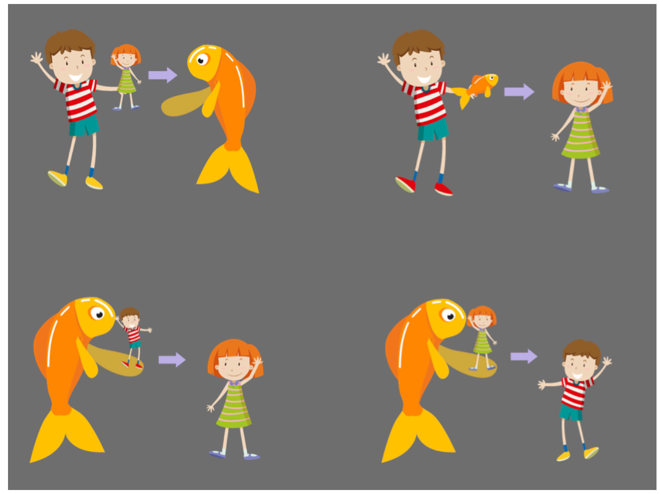

Fig. 1. The visual stimuli in the 4AFC test.

Visual stimuli For each story, four digital color images depicted the three story characters in four different agent-object-recipient relations to each other (Fig. 1). Each image featured an arrow intended to make the direction of the action (e.g., who gave what to whom) more explicit. For each version of each story, only one image corresponded to the central event described in the story and was therefore the correct choice. For instance, the correct match for the target sentence (S3) *the boy bought a goldfish for his sister* would be the top-right image in Fig. 1. The three remaining pictures were foils corresponding to possible misinterpretations of the narrative. These foils were designed to depict misinterpretations that were likely to be elicited by three different processing biases:

(i) *Pragmatic context bias*: an incorrect interpretation of the target sentence driven by the expectations set in the pragmatic context of the story (S1+S2). For instance, given the following pragmatic context: *The goldfish walked into the pet store. His younger sister had been wanting a boy for a long time, and now it was time for her to get one*, and the following target sentence: *The boy bought a goldfish for his sister*, a pragmatic-context bias would be indicated by the participant picking the bottom-left image in Fig. 1, instead of the correct picture match (the top-right image);
(ii) *Real-world semantics bias*: an incorrect interpretation of the narrative in which the target sentence is misinterpreted to match what is plausible in the real world. For instance, given the target sentence *The goldfish bought a boy for his sister*, choosing the top-right image in Fig. 1 (instead of the correct bottom-left image) would indicate a real-world semantic plausibility bias;
(iii) *Syntactic bias*: an incorrect interpretation of the narrative in which marked target-sentence syntax is misinterpreted as unmarked syntax (e.g., the double object construction is misread as prepositional object one), or vice versa. For instance, misinterpreting the target sentence *The boy*

bought the sister the goldfish as *The boy bought the sister for the goldfish* (through the accidental insertion of the preposition *for*) would result in the participant mistakenly clicking on the incorrect top-left image, instead of the correct top-right image.

Given the different narrative structure of each story, a one-to-one mapping between the three picture foils and the three processing biases under scrutiny was not achievable in every trial. However, we estimated that the chances of identifying the three biases in incorrect choices would be equally high when looking across all trials from each participant.

Procedure

Participants sat in front of a computer screen and wore headphones for the entire procedure. Responses in the 4AFC tests were given with a mouse click. Instructions were presented on screen in Norwegian Bokmål and were identical for all participants; however, the participants in the Noise group were advised orally about the presence of noise in the stimuli. The experiment was programmed in PsychoPy2 version 1.90.3 (Peirce & MacAskill, 2018) and began with a practice story (with plausible pragmatic context, plausible target-sentence semantics, and unmarked target-sentence syntax) intended to familiarize the participants with the procedure. After familiarization, the eight stories were presented in fully randomized order. Each story was introduced by a 3 s countdown on screen, after which the sound file was played and a drawing of the three characters of the story were shown on screen (order of presentation for the three characters was fully randomized across participants). After the end of the story, four pictures were presented at the four corners of the screen (as shown in Fig. 1), and the participants were instructed to click at the picture corresponding to what they thought to be the main event in the story. Mouse cursor position was reset at the center of the screen for each 4AFC test.

Data analysis

Accuracy and response time (RT) data were recorded by the experiment script. All possible types of incorrect responses were manually coded as being either due to a pragmatic context bias, a real-world semantics bias, a syntactic bias, or to a combination of two or more biases (for cases in which the incorrect choices were likely to be due to multiple biases). Data pre-processing and statistical analyses were run using R version 3.5.0 (R Core Team, 2018) in RStudio 1.2.1186. Linear mixed-effects models were run using the package lme4 version 1.1-19 (Bates, Maechler, Bolker, & Walker, 2015) and lmerTest 3.0-1 (Kuznetsova, Brockhoff, & Christensen, 2017). All accuracy (correct vs. incorrect) models were logistic mixed-effects models fit through maximum likelihood (Laplace Approximation) with a BOBYQA-optimizer. In addition to accuracy, we analyzed RTs for accurate answers using linear mixed-effects models with log-rescaled outcome variable. All models included random intercepts for subjects and items (random slopes were omitted for model convergence reasons). In the case of null results, we ran Bayes Factor analyses to get indication of whether there was evidence in favor of the null hypothesis, using the brms package (Bürkner, 2017) in R. All Bayesian models had weakly conservative priors for intercept (normal[μ=0, σ=1]), beta estimates (normal[μ=0, σ=1]), SDs of random effects (normal[μ=0, σ=.2]), as well as for correlation coefficients in interaction models (lkj[η=5]).

Results

Accuracy and RTs

To map the relative weight of pragmatic, semantic, and syntactic information sources in noisy and noise-free conditions, we looked at accuracy, response time (RT), and rate and types of errors. For both the No-noise group and the Noise group, overall accuracy on the 4AFC test was high. The average proportion of trials in which participants clicked on the correct picture was 0.78 (within-subject SD = 0.25) in the No-noise group, and 0.69 (within-subject SD = 0.21) in the Noise group. This difference was statistically significant (*Correct ~ Noise + ε*: β = –0.92, *SD* = 0.41, z = –2.25, p = .024), suggesting an overall detrimental effect of perceptual noise on comprehension. No statistically significant difference in RTs was found across conditions (*RTs ~ Noise + ε*: β = 0.38, *SE* = 0.69, t = 0.55, p = .58). We found no cumulative main effect of semantic plausibility and syntactic markedness on accuracy (*Correct ~ Plausibility/Markedness + ε*: β = –0.53, *SD* = 0.14, z = –0.38, p = .7) and RTs (*RT ~ Plausibility/Markedness + ε*: β = 0.01, *SE* = 0.32, t = 0.45, p = .65). A Bayes Factor analysis indicated substantial evidence for the null hypothesis (BF = 28.51, Post.Prob. = 0.97), suggesting that the concurrence of semantic implausibility and syntactic markedness did not consistently result in worse performance, compared to stories with plausible content and unmarked syntax. However, when looking at the three information sources individually, a significant main effect of syntactic markedness was found on accuracy (β = –1.5, *SD* = 0.36, z = –4.14, p < .001), revealing ca. 18% lower accuracy for target sentences with marked syntactic structures (i.e., double-object). We also found a statistically significant main effect of story-internal congruence on accuracy (*Correct ~ Congruence + ε*: β = –3.45, *SD* = 0.56, z = –6.11, p < .001) and RTs (*RTs ~ Congruence + ε*: β = 0.29, *SE* = 0.06, t = 4.74, p < .0001): accuracy was higher and RTs faster for stories in which the events described in S1+S2 and S3 were congruent with each other, and irrespective of whether the two cues were both plausible or implausible (*Correct ~ Congruence × Plausibility + ε*: β = 0.04, *SD* = 0.45, z = 0.09, p = .92) and RTs (*RTs ~ Congruence × Plausibility + ε*: β = 1.1, *SE* =

0.61, $t = 1.79$, $p = .076$).[1] Moreover, the effect of congruence was independent of the main effect of syntactic markedness observed above (accuracy, *Correct ~ Congruence × Syntax + ε*: $\beta = -0.04$, $SD = 1.62$, $z = -0.07$, $p = .94$; RTs, *RTs ~ Congruence × Syntax + ε*: $\beta = 0.15$, $SE = 0.82$, $t = -0.18$, $p = .85$). However, a Bayes Factor analysis did not provide substantial evidence for the null hypothesis in this case, suggesting that additional data is needed (BF = 1.11, Post.Prob. = 0.52).

Error analysis

In order to individuate how the three information sources were weighted during processing, and how they might be driving comprehension errors, we performed an error analysis. For this purpose, we looked at incorrect responses in situations of story-internal incongruence only, since pragmatic and semantic bias can only be fully distinguished in this case. Distribution of errors is presented in Fig. 2. Across conditions, pragmatics-biased errors accounted for 54% of all errors (No-noise = 22% (42 errors), Noise = 32% (97 errors)); semantics-biased errors accounted for 26% (No-noise = 8% (14 errors), Noise = 18% (55 errors)); and syntax-biased errors accounted for 20% (No-Noise = 8% (15 errors), Noise = 12% (36 errors)). Both semantic bias ($\beta = 0.94$, $SE = 0.04$, $t = 2.02$, $p = .043$) and pragmatic bias ($\beta = 0.46$, $SE = 0.04$, $t = 9.9$, $p < .001$) drove significantly more incorrect responses than syntactic bias; syntactic bias was in turn significantly different from zero ($\beta = 0.26$, $SE = 0.034$, $t = 7.79$, $p < .001$, model structure: *Response ~ Bias + ε*). We found no significant two-way interactions between the three sources of bias taken individually (i.e., pragmatics, semantics, and syntax) and noise, suggesting that the role of these information sources in eliciting incorrect responses was not affected selectively by the presence of noise. However, Fig. 3 indicates an evident increase in responses due to a semantic bias, when noise was added to the input, although this interaction was not significant: $\beta = 0.16$, $SE = 0.1$, $t = 1.6$, $p = .11$. A Bayes Factor analysis did not provide robust evidence for this null result (*Noise × Semantics + ε*: BF = 1.63, Post.Prob. = 0.62), suggesting that further investigation is needed.

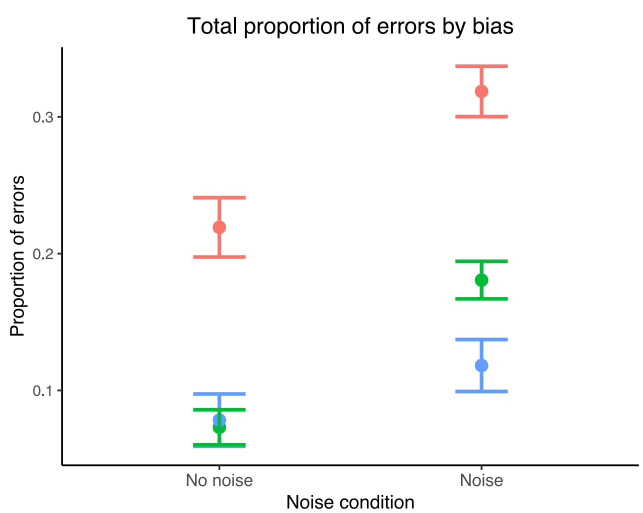

Fig. 2. Distribution of information source biases in incorrect responses (incongruent trials only)

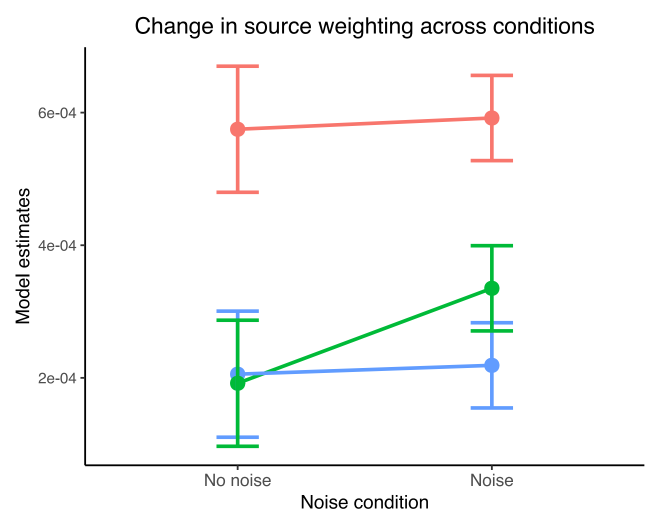

Fig. 3. Predicted values for the model *Response ~ Bias × Noise + ε*

Discussion

In this initial study, we investigated how three sources of information commonly acknowledged in the literature on linguistic processing (i.e., pragmatic/contextual expectations, real-world semantics, and syntactic structure) might contribute differently and dynamically to listeners' comprehension of spoken language input in noisy vs. no-noise conditions. Participants were presented with short stories, in which the three information sources under scrutiny either pointed unequivocally to the same interpretation of the narrative or toward conflicting interpretations. This allowed us to assess the relative weight listeners allocated to the different kinds of information in their interpretation of the linguistic input. Half of the participants listened to stories in the presence of Brownian noise. We hypothesized that listeners would change their processing strategy by generally weighting top-down information more in situations of perceptual noise than in noise-free situations. Moreover, we asked whether the relative weight given to the individual information sources would change when noise was added.

The results provided initial support for our hypothesis by showing that listeners relied more on top-down information in noisy contexts, compared to noise-free ones. In general, accuracy was lower for the Noise group, reflecting the fact that the presence of perceptual noise impedes processing. In both Noise and No-noise groups, listeners made incorrect responses that reflected processing biases driven by either the pragmatic, semantic, or syntactic information in the input—

[1] In the models, plausibility was coded as –1 (S1+S2 and S3 = implausible), 1 (S1+S2 = plausible, S3 = implausible), 2 (S1+S2 = implausible, S3 = plausible), and 3 (S1+S2 and S3 = plausible).

though this happened almost twice as often in the Noise group compared to the No-noise group. Moreover, we found indications that the relative weighting of the different information cues may change when noise is added, with real-world semantics gaining more weight. A number of computational models of language comprehension (e.g., Frank, Koppen, Noordman, & Vonk, 2003, 2008; Venhuizen et al., 2019) have shown that integrating knowledge about the world with lower-level representations of the linguistic input leads to more accurate inferences about the intended meaning of the input. It is possible that the presence of perceptual noise in the signal pressures the processing system and makes it harder for the listener to establish solid representations of the incoming input (e.g., of its syntactic structure and of its pragmatic/contextual information): this may push the listener to rely more on knowledge that is stable over time (i.e., semantic knowledge of the world; see e.g., Kintsch, Patel, & Ericsson, 1999). This mechanism would explain the increase in errors driven by a real-world semantics bias in conditions of noisy signal, but not of those driven by syntax and pragmatics (which are more dependent on establishing representations of the incoming input on the fly). However, this result is only tentative and will need further investigation with more statistical power. Note also that our experimental design only allowed to test comprehension offline (by allowing the participants to make a choice after the end of the story), therefore increasing memory pressure. A more online version of the paradigm (e.g., one that uses mouse tracking/eye tracking) may provide further insights into this issue.

Other interesting results emerged from the study. First, we found a significant main effect of congruence between the pragmatic context of the story and the semantics of the target sentence, with both noisy and non-noisy stimuli. This can be explained in terms of the previously observed mutual influence between story-internal coherence and semantics-based inferences in language comprehension (see e.g., Frank et al., 2003). Second, we found that whenever the pragmatic context of the story and the target-sentence semantics were incongruent (e.g., *the boy walked into the pet store → the goldfish bought a boy for its sister*), the pragmatic context "attracted" the listeners' incorrect interpretations to a significantly larger extent than real-world semantics. This evidence is in line with, for instance, previous ERP evidence from Nieuwland and Van Berkum (2006), who showed that listeners' natural tendency to assume animate characters (in our case, human-animate vs. nonhuman-animate) as being agents in stories can be overruled by counterfactual discourse contexts. Third, we found a significant main effect of syntax markedness in the target sentence (S3), in both noisy and noise-free situations, revealing that sentences with a double-object structure are consistently associated with lower accuracy, than sentences with prepositional dative structure. This finding adds to previous psycholinguistic literature documenting the effects of syntactic markedness on language processing (Dabrowska & Street, 2006), and nicely replicates the results from Gibson et al. (2013) and Gibson et al. (2017), in which prepositional dative sentences were shown to lead to literal (although semantically implausible) readings of the sentences more often compared to double-object sentences.

Existing models of language processing under conditions of acoustic challenge (e.g., in hearing-impaired populations) propose that listeners compensate for degraded input by increasing their cognitive effort in terms of memory, attention-based performance monitoring, and allocation of (extralinguistic) neurocognitive resources (e.g., Eckert, Teubner-Rhodes, & Vaden, 2016; Peelle, 2018). However, these compensatory top-down mechanisms have traditionally been thought to only become relevant as a "last resort", when all bottom-up information fails. Instead, our results may suggest that top-down information critically contributes to language processing by default—and more so when the signal itself becomes degraded and therefore less informative. Moreover, our findings hint at a hierarchical weighting of information sources that is flexibly changed in noisy processing situations—at least when the language input is internally incongruent (see e.g., Yurovsky et al., 2017). Reliance on top-down pragmatic context and real-world semantics is largely increased when the language input is degraded by perceptual noise: listeners may rely more heavily on top-down strategies to compensate for the reduced informativeness of the bottom-up cues. Priorities for future studies using the sentence-picture matching design presented here include focusing on languages other than Norwegian, as well as on cross-linguistic differences in the weighting of top-down information. Moreover, it may be important to move away from a binary noise vs. no-noise manipulation and toward a more continuous variation of the amount of noise added to the signal. This may not only lead to stronger patterns of results but also give rise to interesting non-linearities in the data.

Conclusions

Successful language processing depends on the seamless and rapid integration of bottom-up and top-down information. When the bottom-up signal is degraded by noise (as it happens in many everyday situations), listeners become more reliant on top-down information sources. This study presents a novel methodological framework within which to investigate the simultaneous contribution and dynamic weighting of three top-down information sources—pragmatic context, real-world semantics, and sentence syntax—to language processing in the presence of perceptual noise. Our results nicely dovetail with previous findings, while highlighting the need for a unified model of the relative weighting of bottom-up and top-down information in spoken language processing in noisy situations.

Acknowledgments

This research was supported by the Danish Council for Independent Research (FKK) Grant DFF-7013-00074 awarded to Morten H. Christiansen. We are grateful to three anonymous reviewers for useful comments and suggestions for improvement.

References

Bates, E., & MacWhinney, B. (1989). Functionalism and the Competition Model. In B. MacWhinney and E. Bates (Eds.), *The Crosslinguistic Study of Sentence Processing*, 3–73. Cambridge: Cambridge University Press.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48.

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), 1-28.

Christiansen, M.H. & Chater, N. (2016). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral & Brain Sciences, 39,* e62.

Dąbrowska, E., & Street, J. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences*, *28*, 604-615.

Eckert, M. A., Teubner-Rhodes, S., & Vaden, K. I. (2016). Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions. *Ear & Hearing*, *37*(Suppl. 1), 101S-110S.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164-203.

Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science*, *27*(6), 443-448.

Fillenbaum, S. (1974). Pragmatic normalization: Further results for some conjunctive and disjunctive sentences. *Journal of Experimental Psychology*, *102*, 574–578.

Frank, S. L., Koppen, M., Noordman, L. G. M., & Vonk, W. (2003). Modeling knowledge-based inferences in story comprehension. *Cognitive Science*, *27*, 875-910.

Frank, S. L., Koppen, M., Noordman, L. G. M., & Vonk, W. (2008). World knowledge in computational models of discourse comprehension. *Discourse Processes*, *45*(6), 429-463.

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences*, *110*, 8051–8056.

Gibson, E., Tan, C., Futrell, R., Mahowald, K., Konieczny, L., Hemforth, B., & Fedorenko, E. (2017). Don't underestimate the benefits of being misunderstood. *Psychological Science*, *28*(6), 703-712.

Kintsch, W., Patel, V. L., & Ericsson, K. A. (1999). The role of long-term working memory in text comprehension. *Psychologia*, *42*, 186-198.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26.

Larsen, E. A., & Johansson, C. (2008). Animacy and canonical word order — Evidence from human processing of anaphora. In C. Johansson (Ed.), *Proceedings of the Second Workshop of Anaphora Resolution*, 55-61. Tartu, Estonia: Tartu University Library.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126–1177.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71-102.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1-86.

Martin, A. E. (2016). Language processing as cue integration: Grounding the psychology of language in perception and neurophysiology. *Frontiers in Psychology*, *7*(120), 1-17.

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*, 953–978.

Nieuwland, M. S., & Van Berkum, J. J. A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, *18*(7), 1098-1111.

Peelle, J. E. (2018). Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear & Hearing*, *39*, 204-214.

Peirce, J.W., & MacAskill, M.R. (2018). *Building Experiments in PsychoPy*, London: SAGE

R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell Systems Technical Journal*, *XXVII*, 379–423.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, *49*, 238-299.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, *45*, 447–481.

Szewczyk, J. M., & Schriefers, H. (2001). Is animacy special? ERP correlates of semantic violations and animacy violations in sentence processing. *Brain Research*, *1368*, 108-221.

Venhuizen, N. J., Crocker, M. W., & Brouwer, H. (2019). Expectation-based comprehension: Modeling the interaction of world knowledge and linguistic experience. *Discourse Processes*, *56*(3), 229-255.

Yurovsky, D., Case, S., & Frank, M. (2017). Preschoolers flexibly adapt to linguistic input in a noisy channel. *Psychological Science*, *28*(1), 132-140.

Zwaan, R. A. (2016). Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychonomic Bulletin & Review*, *23*, 1028-1034.