A neurolinguistic investigation into semantic differences of evidentiality and modality[1]

Yurie HARA — *Hokkaido University*
Naho ORITA — *Waseda University*
Deng YING — *The University of Tokyo*
Takeshi KOSHIZUKA — *Tokyo University of Science*
Hiromu SAKAI — *Waseda University*

Abstract. Corpus and EEG studies were conducted to reveal behavioral and neurological differences between processing of Japanese modals and evidentials.

Keywords: evidentiality, modality, causality, electroencephalography, Japanese.

1. Introduction

Many of the prominent studies on inferential evidentials (Izvorski, 1997; Matthewson et al., 2006; McCready and Ogata, 2007) have argued that evidentiality is a kind of modality. Davis and Hara (2014) and Hara (2017) (henceforth, D&H) argue against this evidential-as-modal approach and make two claims: 1. The Japanese sentence-final auxiliary *yooda* 'it seems' as in (1) is a morpheme of evidentiality which is distinct from canonical modal auxiliaries like *daroo* 'I bet/probably' in (2). 2. The semantics of *yooda* relies on the notion of causality.

(1) Ame-ga futta yooda.
rain-NOM fell EVID
'It seems that it rained.'

(2) Michi-ga nureteiru daroo.
streets-NOM wet I-bet
'The streets are wet, I bet/probably.'

The goal of this paper is to justify D&H's theoretical claims by way of corpus and EEG experiment. This paper is structured as follows: Section 2 reviews D&H's argument and Hara's (2017) formal analysis. In particular, *yooda* and *daroo* belong to distinct categories and the semantics of *yooda* presupposes a causal relation and asserts that the speaker perceives the effect state of the causal relation. Crucially, the assertional content of *yooda* does not involve semantics of modality, i.e., quantification over possible worlds. In contrast, *daroo* is an epistemic necessity modal and does not have a causality presupposition. Section 3 presents the research questions of the study. The current study investigates whether the corpus and EEG experiment support the idea that evidentialy and modality are separate categories. Section 4 discusses the result of the corpus study, which confirms the native speaker's intuition that *p* in *p-yooda* expresses a cause event. *Yooda* tends to follow cause-denoting predicates, while *daroo* tends to follow state-denoting predicates. The corpus result alone does not reveal the

[1]This project is supported by NINJAL Project "Cognitive Neuroscience of Linguistic Variation in Pragmatic Inference" (PI: Hiromu Sakai), JSPS Kiban (C) "Semantic-Pragmatic Interfaces at Left Periphery: a neuroscientific approach" (18K00589; PI: Yurie Hara), and JSPS Kiban (A) "How does the grammar emerge in the human brain: A view from comparative cognitive neurosciences of East Asian languages" (15H01881; PI: Hiromu Sakai). We are grateful to Aine Ito, Yohei Ozeki, and Masa Yano for their help with analysis and to Tsu-Yin Chen and Saaya Mori for their assistance in data collection.

 In: M. Franke et al. (eds.) *Proceedings of Sinn und Bedeutung 24*, vol. 1, pp.273–290. Osnabrück University.

difference between *yooda* and *daroo* with respect to whether or not the causality requirement is semantically encoded. Section 5 presents our EEG experiment that overcomes this shortcoming of the corpus study. The felicitous *yooda* condition elicits LAN which is an index of a successful causal inference (Baggio et al., 2008; Cohn and Kutas, 2015) (alternatively, the infelicitous condition elicits N400/P600 effect). In contrast, there is no significant difference in the waveforms between the felicitous and infelicitous *daroo* conditions. This suggests that *yooda* encodes the causality requirement in its semantics while *daroo* sentences give rise to a causal inference in pragmatics. Furthermore, the processing of *daroo* elicits N400, which is argued to be an index of an increase of cost in the processing of possible worlds (Dwivedi et al., 2006; Kulakova et al., 2014; Kulakova and Nieuwland, 2016). Section 6 concludes the paper.

2. Modality and Evidentiality

Evidentials are morphemes that signal the source of information a speaker has to make the claim (Aikhenvald, 2004; Willett, 1988). Evidential morphemes are distinguished depending on their evidence types. For instance, Willett (1988) provides the following typology of evidence types.

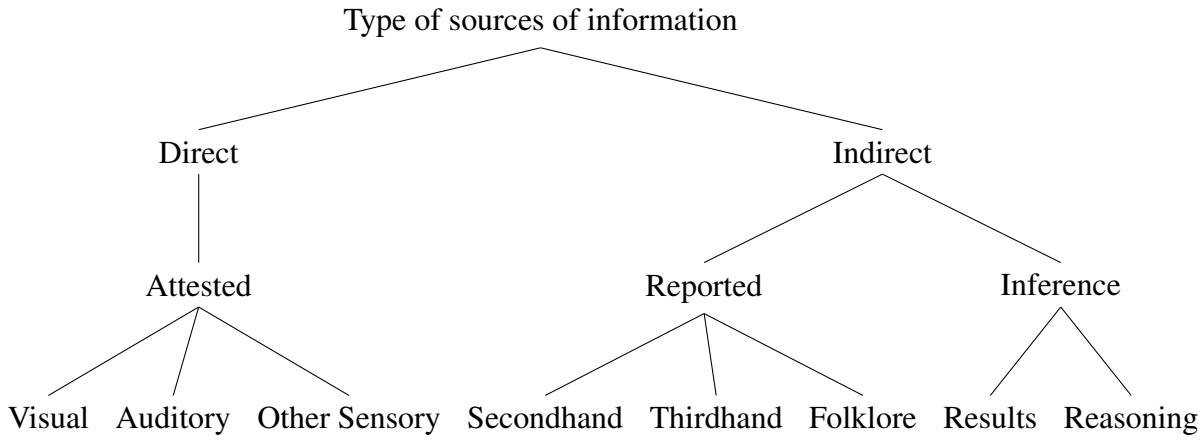


Figure 1: Willett's (1988: 57) typology of evidence types

Japanese has a rich paradigm of indirect evidential markers such as *yooda/mitaida/rasii* 'it seems/appears', *TP+sooda* 'I hear', and *V+sooda* 'looks like' (Aoki, 1986; McCready and Ogata, 2007; Hara, to appear). To our knowledge, the work by McCready and Ogata (2007) is the first to give a formal analysis to the Japanese indirect evidentials. According to McCready and Ogata (2007), evidential markers should be treated as a kind of epistemic modal. Many of the leading literature on evidentiality in other languages (Izvorski, 1997; Matthewson et al., 2006) also classifies evidentiality under the general category of epistemic modality. On the other hand, canonical modals like English *must* are argued to belong to the inferential evidentials (von Fintel and Gillies, 2010; Degen et al., 2019) blurring the boundary between evidentiality and modality. By examining the semantics of Japanese inferential evidential *yooda*, D&H show that evidentials constitute a category distinct from epistemic modality.

2.1. Lack of epistemic commitment

D&H show that *yooda* is different from *daroo* in that the prejacent *p* in *p-yooda* is not an at-issue commitment of the sentence since it is cancellable.

In (3) and (4), both a bare assertion *p* and *p-daroo* commit the speaker to *p*, thus *p* cannot be cancelled.

(3) #Ame-ga futta kedo jitsu-wa futtenai.
rain-NOM fell but actually fall-NEG
'#It rained but in fact it didn't.'

(4) #Ame-ga futta daroo kedo jitsu-wa futtenai.
rain-NOM fell I-bet but actually fall-NEG
'#Probably, it rained but in fact it didn't.'

In contrast, the prejacent *p* in *p-yooda* is cancellable as in (5).

(5) Ame-ga futta yooda kedo, jitsu-wa futte-nai. (Dereka-ga mizu-o maitanda.)
rain-NOM fell EVID but actually fall-NEG someone-NOM water-ACC sprayed
'It seems that it rained, but in fact it didn't. (Someone sprayed water.)'

In short, if *p-yooda* were a kind of modality, it should also give rise to a (weak) commitment to *p*. Since its prejacent is cancellable, *yooda* belongs to a different category.

2.2. Causality

D&H also claim that the notion of evidentiality needs to encode *asymmetric* causal dependencies: Rain causes wet streets but not *vice versa*. McCready and Ogata (2007) propose a Bayesian modal semantics for a number of evidentials. In M&O's analysis, what counts as evidence in asserting *p-yooda* is some information *q* such that learning *q* raised *a*'s subjective probability of *p*. In (6), thus, *a* learns that the streets are wet, which has led *a* to raise her subjective probability of *p*, hence the use of *yooda* is acceptable.

(6) (Looking at wet streets)
Ame-ga futta yooda.
rain-NOM fell EVID
'It seems that it rained.'

However, D&H show that it makes wrong predictions if we switch *p* and *q*, as in (7). Learning that it is raining should also raise the agent's subjective probability of "the streets are wet", thus M&O wrongly predict that *yooda* would be acceptable in (7).

(7) (Looking at falling raindrops)
#Michi-ga nureteiru yooda.
streets-NOM wet EVID
'#It seems that the streets are wet.'

From this observation, D&H propose that what counts as evidence in asserting *p-yooda* is some information *q* such that *p* causes *q*.

Daroo seems to be in a complementary distribution with *yooda*. That is, *p-daroo* is infelicitous when the speaker learns information q such that p causes q as in (8), while it is felicitous when the information q is such that q causes p as in (9).

(8) (Looking at wet streets)
#Ame-ga futta daroo.
rain-NOM fell I-bet
'#Probably, it rained.'

(9) (Looking at falling raindrops)
Michi-ga nureteiru daroo.
streets-NOM wet I-bet
'Probably, the streets are wet.'

We assume with Hara and Davis (2013) that unlike the case of *yooda*, the causality is not lexically encoded in the semantics of *daroo*. Rather, the above contrast is a result of pragmatic competition. *Yooda* has a stronger semantics since it requires a more specific context than *yooda*. Since *yooda* is a better candidate fo the context in (8), *daroo* is infelicitous in (8). See Hara and Davis (2013) for details.

To conclude, the semantics of *p-yooda* is dependent on causality. More specifically, p in *p-yooda* denotes a cause event in a causal relation, p causes q. In contrast, the seeming causal requirement on the use of *daroo* arises as a result of pragmatic competition.

2.3. Semantics of *daroo* and *yooda*

Given D&H's observation, Hara (2017) formulates the semantics of *yooda* in Kaufmann's (2013) Causal Premise Semantics, which incorporates causal networks to Kratzer's (2005) premise semantics. Let $\mathbf{f}$ and $\mathbf{g}$ be premise background and ordering source respectively and w a possible world. Premise structure $\text{Prem}((\mathbf{f}^*\mathbf{g})(w))$ is obtained by ranking Kratzer premise sets.[2] Basic modal operators, Must and May, are defined as follows:

(10) a. $\text{Must}(p)$ is true at $\mathbf{f},\mathbf{g},w$ iff p is a necessity relative to $\text{Prem}((\mathbf{f}^*\mathbf{g})(w))$.
b. $\text{May}(p)$ is true at $\mathbf{f},\mathbf{g},w$ iff p is a possibility relative to $\text{Prem}((\mathbf{f}^*\mathbf{g})(w))$.

We propose that *daroo* is a necessity modal, with the following interpretation:[3]

(11) **Interpretation of *daroo*:** $\text{Daroo}(p)$ is true at $\mathbf{f},\mathbf{g},w$ iff $\text{Must}(p)$ is true at $\mathbf{f},\mathbf{g},w$.

Turning to the evidential *yooda*, Hara's (2017) formalization of causality is built on the interpretation of conditionals. In the current framework, we obtain a causal premise background $\mathbf{f}_c[p]$ by hypothetically updating a causal premise background $\mathbf{f}_c$ with the antecedent proposition p as in (12).

(12) **Hypothetical update**: For all w: $\mathbf{f}[p](w) := \{\{p\}\}^*\mathbf{f}(w)$.

Thus, 'p causes q' is formulated as $\text{Must}_p(q)$ in (13):

[2]Many of the definitions are omitted for space reasons. See Hara (2017) and Kaufmann (2013).
[3]*Daroo* has some other interesting lexical restrictions which are irrelevant to the current purpose of the study. See Hara (2006, 2018).

(13) **Causal rule**: $\text{Must}_p(q)$ is true at $\mathbf{f}_c, \mathbf{g}, w$ iff $\text{Must}(q)$ is true at $\mathbf{f}_c[p], \mathbf{g}, w$.

Finally, we define the interpretation of *yooda*. As shown by D&H, asserting *p-yooda* does not commit the speaker to p, thus p contributes to the *antecedent* part of the conditional rather than the consequent, i.e., $\text{Must}_p(q)$. In Hara's (2017) original definition, the causal component $\text{Must}_p(q)$ was part of the asserted content. Here, we differ from Hara's (2017) original analysis and place the causal component in the presupposition since it is part of the common ground and it is not asserted (e.g., in uttering (1), the speaker is not asserting 'rain causes wet streets' but assuming that it is commonly known.). Thus, the interpretation of *yooda* is defined as in (14): $\text{Yooda}(p)$ presupposes that there is some state q such that q is a necessity relative to $\text{Prem}((\mathbf{f}_c[p]^*\mathbf{g})(w))$ and asserts that the speaker perceives q at w.

(14) **Interpretation of *yooda*:**
$\text{Yooda}(p)$ is defined at $\mathbf{f}_c, \mathbf{g}, w$ iff $\exists q$ such that $\text{Must}_p(q)$ is true at $\mathbf{f}_c, \mathbf{g}, w$. (presupposition)
If defined, $\text{Yooda}(p)$ is true at w iff the speaker perceives q at w. (assertion)

As can be seen in (14), the assertion of *p-yooda* is evaluated only at the single actual/utterance world w. As with D&H, our definition (14) correctly derives the interpretations and distributions of *yooda* discussed above. In (5), the speaker presupposes that raining causes wet streets and she only asserts that she observed wet streets, which are usually caused by raining but could be caused by other factors, e.g., someone sprayed water with a hose. (7) causes a presupposition failure since we know, as background knowledge, that wet streets do not cause rain.

In summary, *yooda* belongs to the category of evidentiality which is distinguished from that of modality such as *daroo*. The semantics of *p-yooda* includes a presupposition that there is an event q such that p causes q. The assertion of *p-yooda* does not commit the speaker to p but only entails that the speaker perceives q. In contrast, *p-daroo* is an epistemic necessity modal, the semantics of which involves quantification over possible worlds, thus the assertion of it commits the speaker to p. Furthermore, *daroo* does not have a causality presupposition like *yooda*. As discussed in Section 2.2, the apparent causal restriction that p in *p-daroo* needs to denote an effect state is a result of pragmatic competition.

3. Research Questions

The previous section summarized D&H's claims that evidentials and modals are separate categories and that the interpretation of *yooda* is dependent on causality. The goal of this paper is to empirically justify these claims using corpus and an EEG experiment:

(15) Research Question 1
Can we justify the theoretical claims by way of corpus/EEG studies?

(16) Research Question 2
Are there neural indices that indicate the computational difference between modals and evidentials?

Table 1: NPMI

	yooda		daroo
past	0.22	$>$	0.057
non-past	0.177	$<$	0.214
verb	0.204	$>$	0.188
verb-positive	0.196	$>$	0.17
verb-negative	0.209	$<$	0.253
adjective	0.116	$<$	0.135

4. Corpus

We conducted a corpus study to test whether the distribution of *yooda* is dependent on the causal relation. Since the corpus does not contain information on whether a predicate denotes a cause or effect, we made the assumptions in (17) and the predictions in (18):

(17) a. Causes are events while effects are states.
b. The cause event temporally precedes the effect state.

(18) a. *Yooda* tends to be attached to past-tensed and eventive predicates.
b. *Daroo* tends to be attached to non-past and stative predicates.

To test the predictions in (18), we use Balanced Corpus of Contemporary Written Japanese (BCCWJ) (Maekawa et al., 2014) containing approximately 100 million words collected from various kinds of Japanese texts. We extract sentences ending with *daroo* and *yooda*, resulting in 30686 *yooda* sentences and 47538 *daroo* sentences. We measured Normalized Pointwise Mutual Information (NPMI): The higher positive values of NPMI indicate stronger associations or positive correlations between kinds of predicates and final auxiliaries *yooda/daroo*.[4] Table 1 summarizes the result. As predicted, *yooda* has better associations with past-tensed predicates and positive verbs, which tend to denote events, while *daroo* has better associations with non-past predicates and negative verbs and adjectives, which tend to denote states (Krifka, 1990).

In short, *yooda* is more likely to follow past-tensed predicates and positive verbs, which are likely to denote cause events, than non-past predicates and negative verbs and adjectives, which are likely to denote effect states. This result is in accordance with D&H's claim that causality is crucial in the interpretation of *yooda*: *p* in *p-yooda* denotes a cause event in a causal relation. The current result also shows that *daroo* is more likely to follow non-past predicates, negative verbs and adjectives, which are likely to denote effect states. As discussed in Sections 2.2 and 2.3, the lexical semantics of *daroo* does not involve causality requirement. Thus, we interpret this result as an outcome of the pragmatic competition between *yooda* and *daroo*.

5. EEG

An EEG experiment was conducted to address the research questions laid out above in Section 3.

[4]NPMI is represented by a value between $[-1, +1]$. If two predicates have zero co-occurrence, NPMI is -1. If they are independent, it is 0. If they always co-occur, it is $+1$. See Bouma (2009).

5.1. Participants

Thirty-eight right-handed native Japanese speakers were paid to participate in this study. Six participants were excluded from the analysis because of facility disorder. After artifact rejection, two more participants were excluded from the analysis due to the extremely small number of valid trials in their data (less than 15 trials in at least one of the conditions). The final analysis was conducted with 30 participants (16 males, mean age 20.9 years, age range 18 to 24 years). All participants were right-handed and reported normal vision without any history of neurological or psychiatric disorders. Written consent was obtained from all participants.

5.2. Stimuli

The stimuli had two fully-crossed factors—CONTEXT (Effect-Cause/Cause-Effect) and SFA (sentence-final-auxiliary; *yooda/daroo*)—which resulted in four conditions:

(19) a. EC*y*: Effect-Cause-*yooda* :
Michi-ga nureteiru. Ame-ga futta yooda.
streets-NOM wet rain-NOM fell EVID
'The streets arer wet. It seems that it rained.'

b. CE*y*: Cause-Effect-*yooda*:
#Ame-ga futta. Michi-ga nureteiru yooda.
rain-NOM fell streets-NOM wet EVID
'#It rained. It seems that the streets are wet.'

c. CE*d*: Cause-Effect-*daroo*:
Ame-ga futta. Michi-ga nureteiru daroo.
rain-NOM fell streets-NOM wet I-bet
'It rained. The streets are wet, probably.'

d. EC*d*: Effect-Cause-*daroo*:
#Michi-ga nureteiru. Ame-ga futta daroo.
streets-NOM wet rain-NOM fell I-bet
'The streets are wet. It rained, probably'

EC*y* is a fit condition where the final auxiliary *yooda*, which semantically presupposes that there is a causal relation and attaches to a proposition that denotes the cause event, matches the Effect-Cause context. CE*y* is an anomalous condition where *yooda* is attached to the effect state in a causal relation. CE*d* is a fit condition where the final auxiliary *daroo* is attached to a conclusion drawn from premises. EC*d* is an anomalous condition, but unlike CE*y*, which is semantically anomalous, it is pragmatically anomalous. That is, as argued in Section 2, unlike *yooda*, *daroo* does not lexically encode a causal relation in its semantics. (19d) is anomalous because *yooda*, which semantically encodes the causal requirement, is more optimal in the Effect-Cause context.

All the stimuli were normed for naturalness by Japanese native speakers.

5.3. Procedure

Each condition had 78 items. 78 fillers were included. The experiment was counterbalanced so that one participant will not see the same context twice (234 trials in each experiment). The

procedure of a trial is depicted in Figure 2. Each trial started with a fixation cross shown for 600ms. After 300ms blank screen, each word of the first sentence was presented for 600ms with an inter-word interval of 300ms. After the first sentence was presented, a '~' sign is presented for 600ms to let the participants to blink. After 300ms blank screen, the second sentence was presented. At the end of each filler trial, subjects were asked to give a yes/no response to an acceptability judgement task.

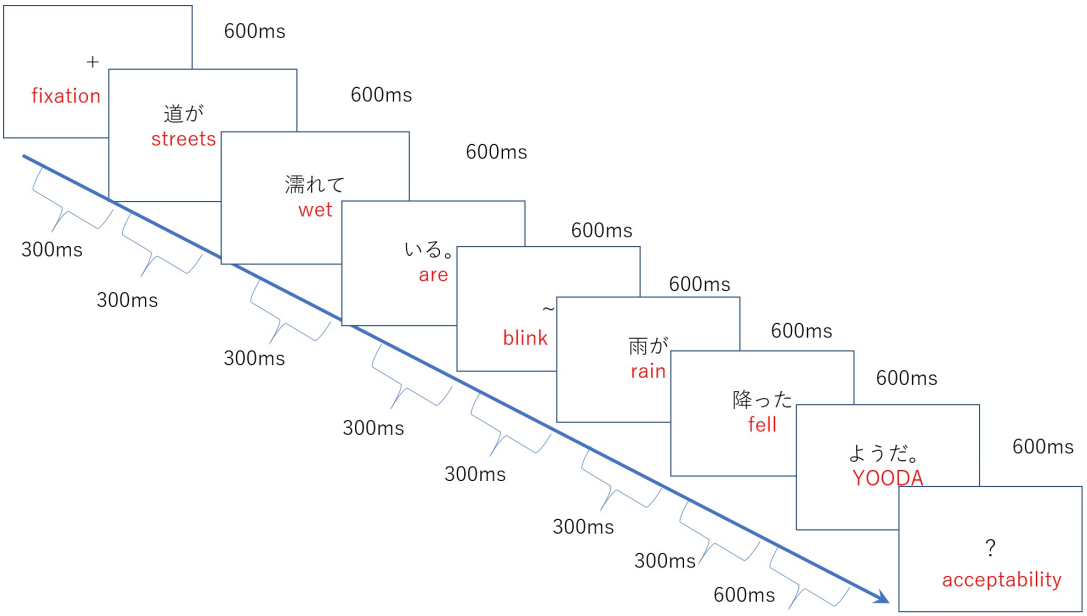


Figure 2: Procedure of a trial

5.4. Recording

EEG was recorded using AC amplifiers (Brain Products) from 32 electrodes equipped in an elastic cap (EASYCAP) according to the International 1020 system. Electrodes included the following positions: FP1, FP2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, C1, Cz, C4, T8, CP5, CP1, CP2, CP6, TP9, P3, Pz, P4, P8, TP10, O1, and O2. Another electrode was placed at nose for online referencing. Offline, all electrode sites were re-referenced to the average of the right and left mastoids. The electro-oculogram (EOG) was located at the outer canthus of right eye in order to detect horizontal eye movements and blinks. Scalp impedances were kept below 5 kΩ. The continuous EEG was digitized at 250 Hz and filtered offline (high-pass at 0.5 Hz and low-pass at 40 Hz).

5.5. Analysis

ERPs were calculated by averaging the EEG time-locked to a point 200ms pre-stimulus onset and lasting until 1000ms after the onset of the stimulus. The 200ms pre-stimulus period was used as a baseline. Trials containing ocular or muscular artifacts, were not taken into consideration in the averaging process. Based on visual inspection of ERP waveforms and the time windows of potential effects, two time windows were selected for statistical analyses: 300-

500ms and 500-700ms. Repeated-measures ANOVA (`anovakun` 4.8.3[5] implemented in R (R Core Team, 2019)) were performed with the factors CONTEXT (Cause-Effect/Effect-Cause), SFA (sentence-final auxiliary; *yooda*/*daroo*) and ROI (regions of interests). Six ROIs are obtained by calculating the mean of the three electrodes in each region (anterior-left (AL; F3 ,F7, FC5); anterior-right (AR; F4, F8, FC6); central-left (CL; FC1,CP1,CP5); central-right (CR; FC2,CP2,CP6); posterior-left (PL; P3, P7, O1); posterior-right (PR; P4, P8, O2)) as in Figure 3.

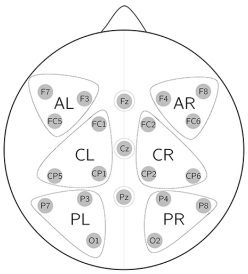
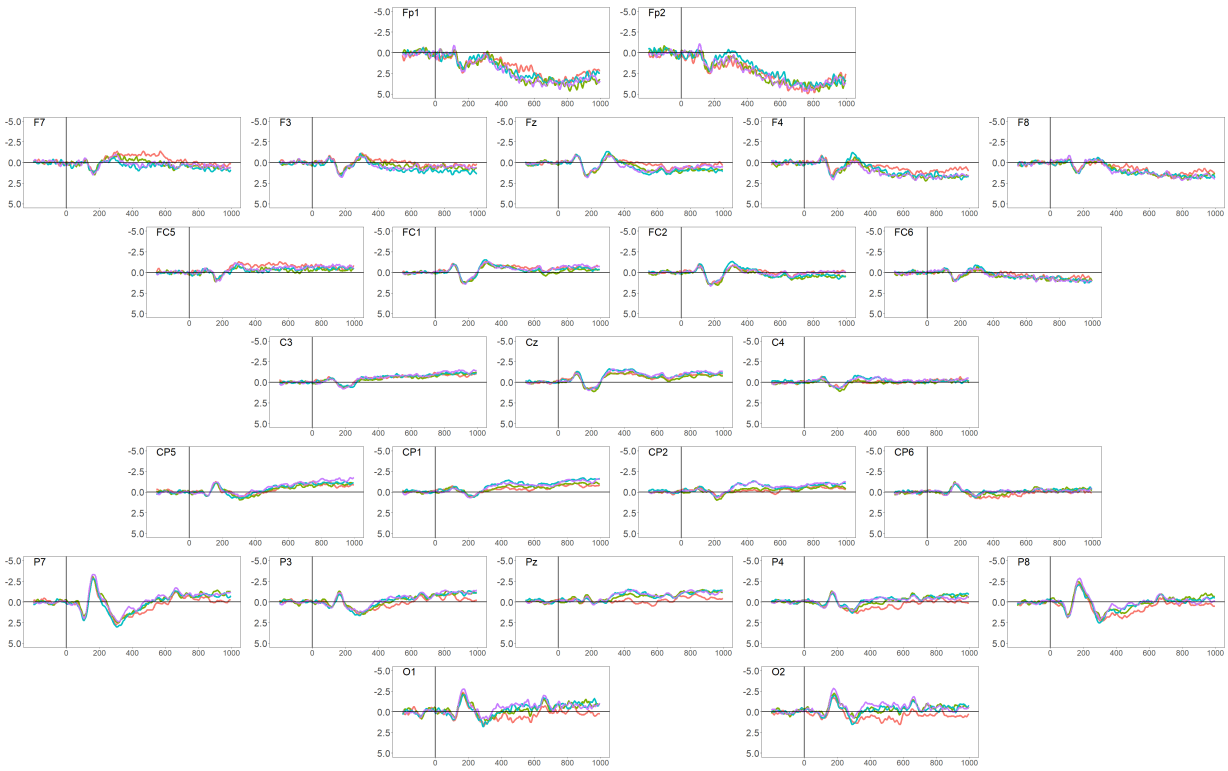


Figure 3: 6ROI (Taken from Kulakova et al. (2014))

5.6. Result

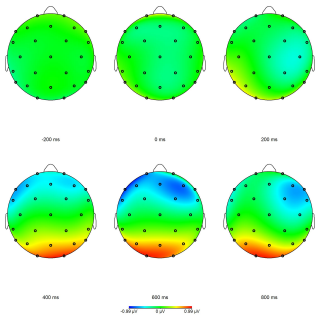
A $2 \times 2 \times 6$ repeated measures ANOVA with the factors CONTEXT (CE/EC), SFA (*y*/*d*) and ROI in the 500-700ms time window revealed a three-way interaction ($F(5,145) = 3.19, p < .01$; Figure 4). The EC*y* condition elicited a significantly more negative-deflection at AL compared to CE*y* ($F(1,29) = 8.06, p < .01$). At PR, the CE*y* condition elicited a significantly more negative-deflection at ($F(1,29) = 7.03, p < .05$) compared to EC*y*. In EC conditions, *yooda* was significantly more negative compared to *daroo* in AL ($F(1,29) = 18.63, p < .001$), PL ($F(1,29) = 5.63, p < .05$) and PR ($F(1,29) = 17.51, p < .001$). There was no significant difference between EC*d* and CE*d* conditions nor between CE*y* and CE*d*. Detailed test-statistics in the 500-700ms time window are shown in Table 2.

[5] `http://riseki.php.xdomain.jp/`



(a) Grand average waveforms time-locked to the final auxiliary

(b) Conditions



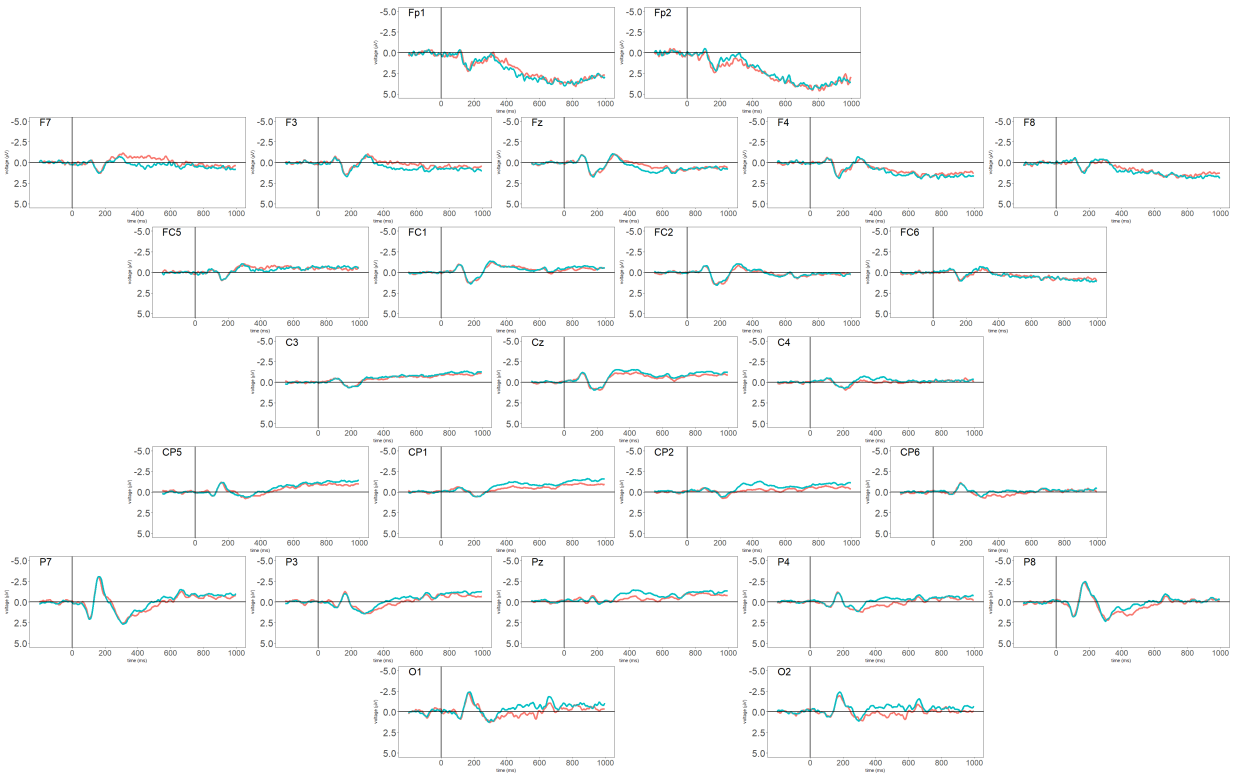
(c) ECy−CEy

Figure 4: 3-way interaction

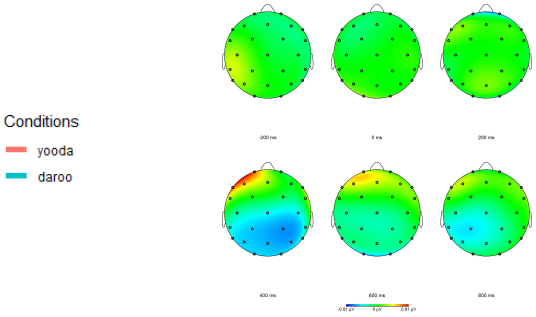
	df	*F*	*p*	sig.
SFA	1,29	3.78	0.06	+
CONTEXT	1,29	0.98	0.32	ns
ROI	5,145	7.11	< 0.001	***
SFA×CONTEXT	1,29	0.07	0.786	ns
SFA×ROI	5,145	4.95	< 0.001	***
CONTEXT×ROI	5,145	1.68	0.1409	ns
SFA×CONTEXT×ROI	5,145	3.19	0.009	**
SFA(*yooda*)				
CONTEXT×ROI	5,145	4.09	0.0017	**
CONTEXT at AL	1,29	9.57	0.004	**
CONTEXT at AR	1,29	2.08	0.15	ns
CONTEXT at CL	1,29	0.35	0.55	ns
CONTEXT at CR	1,29	0.09	0.75	ns
CONTEXT at PL	1,29	2.95	0.09	+
CONTEXT at PR	1,29	6.88	0.01	*
SFA(*daroo*)				
CONTEXT×ROI	5,145	0.18	0.96	ns
CONTEXT(EC)				
SFA×ROI	5,145	7.2692	< 0.001	***
SFA at AL	1,29	18.63	< 0.001	***
SFA at AR	1,29	1.13	0.29	ns
SFA at CL	1,29	0.14	0.70	ns
SFA at CR	1,29	0.57	0.45	ns
SFA at PL	1,29	5.63	0.02	*
SFA at PR	1,29	17.51	< 0.001	***
CONTEXT(CE)				
SFA×ROI	5, 145	0.40	0.84	ns

Table 2: Test-statistics of the ERPs in the 500-700ms time window

A 2×6 repeated measures ANOVA with the factors SFA and ROI in the 300-500ms time window revealed that there was a significant SFA × ROI interaction ($F(5,145) = 7.12, p < .001$; Figure 5). *Daroo* elicited a significantly more positive-deflection at AL ($F(1,29) = 12.44, p < .01$) and a significantly more negative-deflection at central (CL: $F(1,29) = 5.30, p < .05$; CR: $F(1,29) = 12.97, p < .01$) and posterior (PL: $F(1,29) = 4.85, p < .05$; PR: $F(1,29) = 12.14, p < .01$) regions compared to *yooda*. There was also a significant CONTEXT × ROI interaction ($F(5,145) = 3.00, p < .05$). The CE conditions were significantly more negative compared to the EC conditions at PL ($F(1,29) = 4.44, p < .05$) and PR ($F(1,29) = 6.05, p < .05$). Detailed test-statistics in the 300-500ms time window are shown in Table 3.



(a) Grand average waveforms time-locked to the final auxiliary



(b) Conditions (c) *daroo*−*yooda*

Figure 5: SFA × ROI interaction

	df	*F*	*p*	sig.
SFA	1,29	14.99	< 0.001	***
CONTEXT	1,29	2.78	0.10	ns
ROI	5,145	5.11	< 0.001	***
SFA×CONTEXT	1,29	0.006	0.93	ns
SFA×ROI	5,145	7.12	< 0.001	***
CONTEXT×ROI	5,145	3.00	0.01	*
SFA×CONTEXT×ROI	5,145	1.63	0.15	ns
SFA×ROI				
SFA at AL	1,29	12.44	0.001	**
SFA at AR	1,29	0.58	0.45	ns
SFA at CL	1,29	5.30	0.02	*
SFA at CR	1,29	12.97	0.001	**
SFA at PL	1,29	4.85	0.03	*
SFA at PR	1,29	12.14	0.001	**
CONTEXT×ROI				
CONTEXT at AL	1,29	0.79	0.37	ns
CONTEXT at AR	1,29	1.74	0.19	ns
CONTEXT at CL	1,29	0.15	0.69	ns
CONTEXT at CR	1,29	0.006	0.93	ns
CONTEXT at PL	1,29	4.44	0.04	*
CONTEXT at PR	1,29	6.05	0.02	*

Table 3: Test-statistics of the ERPs in the 300-500ms time window

5.7. Discussion

5.7.1. Semantics/Pragmatics of Causality

We interpret the anterior negativity elicited by the EC*y* condition (Figure 6) as Left-Anterior Negativity (LAN), which is argued to be a correlate of successful causal inference (Baggio et al., 2008; Cohn and Kutas, 2015).

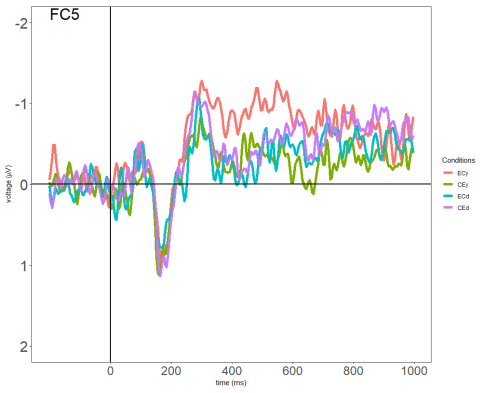


Figure 6: LAN (elicited by EC*y*) or P600 (elicited by CE*y*)

An alternative interpretation is to regard an anterior negativity elicited by the EC*y* condition (Figure 6) as a *positivity* elicited by the CE*y* condition. In this interpretation, the negativity observed in Figure 7 and the positivity observed in Figure 6 are N400/P600 pattern caused by semantic anomaly and reanalysis. Frontal P600 may seem unconventional because P600 elicited by grammatically anomalous constructions is usually observed in posterior regions (e.g., Coulson et al., 1998). However, Kaan and Swaab (2003) report that frontal P600 is an index of ambiguity resolution and processing difficulty in discourse. Indeed, Dwivedi et al. (2006) observe left-frontal P600 for anomalous discourse. Furthermore, when the experiment involves an acceptability judgment task, it is reported that semantically anomalous constructions tend to yield N400 followed by P600 (Kuperberg, 2007; Kulakova et al., 2014), although in Kuperberg (2007); Kulakova et al. (2014), P600 is found in central/posterior regions.

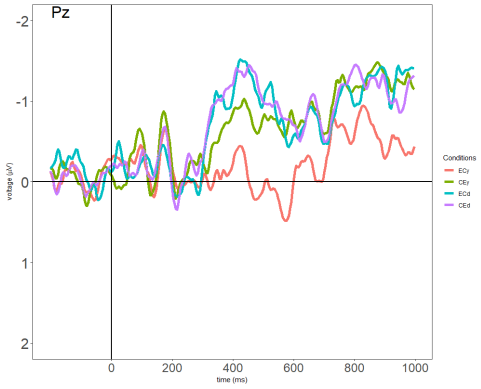


Figure 7: N400 (elicted by CE*y*)

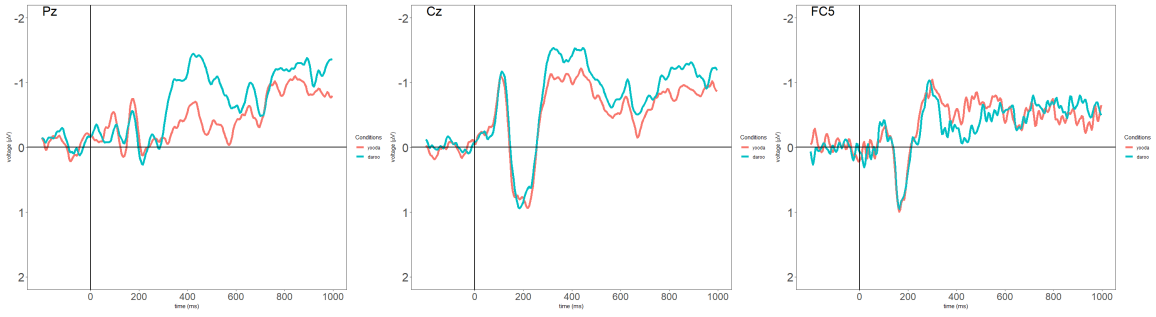
Between the two interpretations, we find the first more tenable than the latter. If the EC*y* condition elicits LAN, only EC*y* is predicted to deflect negative compared to the other three conditions. Similarly, if CE*y* elicits P600, only CE*y* is predicted to deflect positive compared to the other three. In the current experiment, there was a significant difference not only between EC*y* and CE*y* but also between EC*y* and EC*d*. That is, EC*y* was more negative compared to both CE*y* and EC*d* at AL. In contrast, there was no significant difference between CE*y* and CE*d*. Thus, we regard the negativity elicited by EC*y* is LAN, which indexes successful causal inference.

While significant differences between the two *yooda* conditions were obtained, there was no significant difference between the two *daroo* conditions, CE*d* and EC*d*. We attribute this difference between two auxiliaries to the status of causality in each morpheme. That is, as argued in Section 2, the EC*d* is only pragmatically anomalous since there is a better auxiliary, namely *yooda*, which fits the Effect-Cause context. Thus, pragmatic anomaly is not strong enough to affect the waveforms. In contrast, CE*y* is semantically anomalous since the semantics of *yooda* requires the morpheme to be attached to the cause event of a cause-effect relation.

5.7.2. Modality and N400

We regard the posterior and central negativity observed as SFA × ROI interaction for *daroo* (Figure 8) as N400. Let us consider the frequencies of the two morphemes, *yooda* and *daroo*, because we compare them directly and the amplitude of the N400 is known to be affected by the frequency of words (Van Petten and Kutas, 1990). That is, infrequent morphemes elicit amplified N400 compared to high frequent ones. As seen in Section 4, however, *daroo* is more frequent than *yooda* in BCCWJ (30686 *yooda* sentences and 47538 *daroo* sentences). Since *daroo*, which is more frequent than *yooda* amplifies N400, the observed N400 cannot be explained in terms of the frequencies of the morphemes. Instead, we propose that the observed N400 is related to the processing of modality which involves processing of multiple possible worlds that poses increased processing demands. This accords well with the findings in the previous studies: Processing of English and German counterfactuals and modals elicit negativity, which is analyzed as an indication of increased processing costs. Dwivedi et al. (2006) report that hypothetical contexts in English elicit negative trend. Kulakova et al. (2014) observe that German subjunctives elicit LAN compared to indicatives. Kulakova and Nieuwland (2016) show that English counterfactuals elicit larger N400 compared to indicatives. We do not have much to say about the frontal positivity (FC5 in Figure 8(c)). Since this positivity is observed in the same time-window as the posterior negativity is observed, it might also reflect the semantic processing of modality.

An alternative interpretation to the semantic N400 is to interpret it as a result of lexical integration cost. The negativity for *daroo* condition can also be regarded as the positivity for the *yooda* condition. Since *yooda* is more frequently used in the Effect-Cause context as we pointed out in Section 4, its semantic integration might be less costly, which leads to less negativity of the *yooda* condition. This interpretation, however, does not explain why *yooda* in Cause-Effect condition elicits more negativity and cancels out the reduction of negativity in Effect-Cause context. We thus regard the semantic N400 interpretation as more favorable.



(a) Pz (b) Cz (c) FC5

Figure 8: N400 elicited by *daroo*

5.7.3. Summary

We observe that the evidential auxiliary *yooda* which marks an inference from an effect state to a cause event elicited LAN. The alternative interpretation of the result is that the anomalous

CE*y* condition elicits N400/P600 effect of semantic anomaly and reanalysis. On the other hand, the modal auxiliary *daroo* elicits neither effects, even though introspection-based data suggests that EC*d* is anomalous. Our result demonstrates that causal inference involved in *daroo* is different from that of *yooda*. More specifically, EC*d* is pragmatically anomalous while CE*y* is semantically anomalous. Furthermore, our result reveals that processing of modality, which involves processing of multiple possible worlds, induces N400 effect.

6. Conclusion

Evidentiality and modality are very similar notions and often one is subsumed under the other. Davis and Hara (2014), however, argue that evidentiality and modality are separate categories and Hara (2017) offers a formal analysis of the Japanese evidential auxiliary *yooda* that lexically encodes the causal requirement in its semantics. *Yooda* is an evidential morpheme and *p-yooda* presupposes that there is information q such that p causes q and asserts that the speaker perceives q at the actual world $w_@$. Hence, there is no quantification over possible worlds involved in the assertional content of *p-yooda*. In contrast, *daroo* is an epistemic necessity modal which involves universal quantification over possible worlds and its semantics does not encode the causal requirement. To conclude the paper, let us go back to our research questions:

(15) Research Question 1
Can we justify the theoretical claims by way of corpus/EEG studies?

(16) Research Question 2
Are there neural indices that indicate the computational difference between modals and evidentials?

First, the corpus and EEG studies support the idea that the interpretation of *yooda* is dependent on causality and evidentials and modals are separate categories. In the corpus study, *yooda* is more likely to follow cause-denoting predicates than state-denoting predicates. The corpus result also suggests that the interpretation of *daroo* is also dependent on causality since *daroo* is more likely to follow state-denoting predicates than cause-denoting predicates. However, the EEG study shows that the effect of causality is stronger in *yooda* than *daroo*. The EC*y* condition where the causal context matches the semantics of *yooda* elicits LAN (alternatively the anomalous CE condition elicits N400/P600 effect), while no effect is observed for the CE*d* and EC*d* conditions. This result is in accordance with our theoretical claim that *yooda* semantically presupposes a causal relation while *daroo* lacks such a presupposition. The seeming influence of causality in the introspection-based and corpus data is due to a pragmatic competition.

Regarding Research Question 2, we find LAN as an index of successful causal inference in evidentiality and N400 as an index of an increase of processing cost of modality, i.e., multiple possible worlds.

Finally, the study reported in this paper demonstrates that EEG plays a crucial role in dissociating semantic violations from pragmatic ones. Both the introspection-based approach and the corpus study show that the interpretation/distribution of *daroo* is dependent on causality just like *yooda*, but neither can distinguish the nature of the violations.

References

Aikhenvald, A. Y. (2004). *Evidentiality*. Oxford: Oxford University Press.

Aoki, H. (1986). Evidentials in Japanese. In W. Chafe and J. Nichols (Eds.), *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, New Jersey: Ablex Publishing Corporation.

Baggio, G., M. van Lambalgen, and P. Hagoort (2008). Computing and recomputing discourse models: An ERP study. *Journal of Memory and Language 59*, 36–53.

Bouma, G. (2009). Normalized (pointwise) mutual information in collocation extraction. In E. d. C. Chiarcos and Stede (Eds.), *Von der Form zur Bedeutung: Texte automatisch verarbeiten / From Form to Meaning: Processing Texts Automatically, Proceedings of the Biennial GSCL Conference 2009*, Tübingen, pp. 31–40. Gunter Narr Verlag.

Cohn, N. and M. Kutas (2015). Getting a cue before getting a clue: Event-related potentials to inference in visual narrative comprehension. *Neuropsychologia 77*, 267278.

Coulson, S., J. W. King, and M. Kutas (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes 13*, 21–58.

Davis, C. and Y. Hara (2014). Evidentiality as a causal relation: A case study from Japanese *youda*. In C. P. nón (Ed.), *Empirical Issues in Syntax and Semantics 10*.

Degen, J., A. Trotzke, G. Scontras, E. Wittenberg, and N. D. Goodman (2019). Definitely, maybe: A new experimental paradigm for investigating the pragmatics of evidential devices across languages. *Journal of Pragmatics 140*, 33–48.

Dwivedi, V. D., N. A. Phillips, M. Laguë-Beauvais, and S. R. Baum (2006). An electrophysiological study of mood, modal context, and anaphora. *Brain research 1117 1*, 135–53.

Hara, Y. (2006). *Japanese Discourse Items at Interfaces*. Ph. D. thesis, University of Delaware, Newark, DE.

Hara, Y. (2017). Causality and evidentiality. In A. Cremers, T. van Gessel, and F. Roelofsen (Eds.), *Proceedings of the 21st Amsterdam Colloquium*, pp. 295–304.

Hara, Y. (2018). *Daroo* as an entertain modal: an inquisitive approach. In S. Fukuda, M. S. Kim, M.-J. Park, and H. M. Cook (Eds.), *Japanese/Korean Linguistics*, Volume 25. CSLI Publications.

Hara, Y. (To appear). Evidentials: Marking the source of information. In W. M. Jacobsen and Y. Takubo (Eds.), *Handbook of Japanese Semantics and Pragmatics, volume 5 of Handbooks of Japanese Language and Linguistics*. De Gruyter Mouton.

Hara, Y. and C. Davis (2013). *Darou* as a deictic context shifter. In K. Yatsushiro and U. Sauerland (Eds.), *Proceedings of Formal Approaches to Japanese Linguistics 6 (FAJL 6)*, MIT Working Papers in Linguistics 66, pp. 41–56.

Izvorski, R. (1997). The Present Perfect as an Epistemic Modal. *the proceedings of SALT 7*, 222–239.

Kaan, E. and T. Swaab (2003). Repair, revision, and complexity in syntactic analysis: an electrophysiological differentiation. *Journal of Cognitive Neuroscience 15*(1), 98–110.

Kaufmann, S. (2013). Causal premise semantics. *Cognitive Science 37*, 1136–1170.

Kratzer, A. (2005). Constraining premise sets for counterfactuals. *Journal of Semantics 22*, 153–158.

Krifka, M. (1990). Boolean and non-boolean 'and'. In L. Kálmán and P. L. (Eds.), *Papers from the Second Symposium on Logic and Language*, pp. 161188. Budapest, Hungary: Akadémiai Kiadó.

Kulakova, E., D. Freunberger, and D. Roehm (2014, 07). Marking the counterfactual: ERP evidence for pragmatic processing of German subjunctives. *Frontiers in human neuroscience 8*, 548.

Kulakova, E. and M. Nieuwland (2016, 05). Pragmatic skills predict online counterfactual comprehension: Evidence from the N400. *Cognitive Affective & Behavioral Neuroscience in press*.

Kuperberg, G. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain research 1146*, 23–49.

Maekawa, K., M. Yamazaki, T. Ogiso, T. Maruyama, H. Ogura, W. Kashino, H. Koiso, M. Yamaguchi, M. Tanaka, and Y. Den (2014). Balanced corpus of contemporary written japanese. *Language resources and evaluation 48*(2), 345–371.

Matthewson, L., H. Rullmann, and H. Davis (2006). Evidentials are epistemic modals in St'át'imcets. In M. Kiyota, J. L. Thompson, and N. Yamane-Tanaka (Eds.), *Papers for the 41st International Conference on Salish and Neighbouring Languages*, University of British Columbia Working Papers in Linguistics 18, pp. 221–263.

McCready, E. and N. Ogata (2007). Evidentiality, modality and probability. *Linguistics and Philosophy 30*(2), 35–63.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Van Petten, C. and M. Kutas (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition 18*(4), 380–393.

von Fintel, K. and A. S. Gillies (2010). Must . . . stay . . . strong! *Natural Language Semantics 18*(4), 351–383.

Willett, T. (1988). A cross-linguistic survey of the grammaticalization of evidentiality. *Studies in Language 12*(1).