# Linguistic and Social Meaning Match: An experiment on modal concord in English

Mingya Liu & Stephanie Rotter*

**Abstract.** Modal concord (MC) refers to the phenomenon where two modal elements of the same flavor and force in a sentence yield an interpretation of single modality (SM). In this paper, we report on an experimental study on MC in English, addressing their linguistic and social meaning. Our results show a strengthening effect by necessity MC and a weakening effect of possibility MC in that significantly higher speaker commitment ratings were received for necessity MC vs. SM constructions (i.e., *must certainly* vs. *must*) with the reverse pattern for possibility modal constructions (i.e., *may possibly* vs. *may*). Furthermore, MC and SM were shown to differ in social meanings, suggesting a correlation between the meaning strength of a linguistic expression and the social perception of the speaker.

**Keywords.** modal concord; social meaning; experiment; English

**1. Introduction.** Modality is one of the key empirical areas in linguistics, and also one of the most complex ones. This paper deals with the phenomenon 'modal concord' (MC) and addresses the linguistic and social meaning of MC constructions.

As shown in (1), a sentence can be unmodalized (1-a), contain a single modal (SM) element such as a modal auxiliary or adverb (1-b), or include two modal elements including a modal auxiliary and an adverb (1-c).

(1) a. It is the case.
b. It **may** be the case. / It is **possibly** the case. (SM)
c. It **may possibly** be the case. (MC)

In (1-c), the two modal elements are both weak epistemic modals and their co-occurrence can yield the same meaning as just one of them—this phenomenon is labeled as MC in the literature (Geurts & Huitink 2006, Zeijstra 2007, Yalcin 2007), with the discussion dating back to at least the 70s, see an earlier quote and a more recent one below.

> "In most dialects of English not more than one modal verb can occur within the same clause. But both a modal verb and a modal adverb may be combined. When this happens a distinction is to be drawn between modally harmonic* and modally non-harmonic* combinations. For example, 'possibly' and 'may', if each is being used epistemically, are harmonic, in that they both express the same degree of modality, whereas 'certainly' and 'may' are, in this sense, modally non-harmonic. It has been pointed out by Halliday (1970a: 331) that the adverb and the modal verb may, and normally do, "reinforce each other" in a modally harmonic combination; so that, ... or there is a kind of concord running through the clause, which results in the double realization of a single modality." (Lyons 1977: 807-808).

*This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1412, 416591334. Authors: Mingya Liu, Humboldt-Universität zu Berlin (mingya.liu@hu-berlin.de) & Stephanie Rotter, Humboldt-Universität zu Berlin (rotterst@hu-berlin.de).

> "if two modal elements are of the same modal type (epistemic/ deontic/...) and have similar quantificational force (universal/existential), the most salient reading is mostly not a cumulative one but a concord reading, where the semantics seems to contain only one modal operator." (Zeijstra 2007: 137)

With the 'meaning-equivalence' assumption for MC and SM, we can make the inference that the optional adverb *possibly* in (1-c) is semantically vacuous. This is debatable. For example, some authors hold the view that multiple modal expressions have one semantic role but the adverb does make a semantic contribution, giving rise to speaker bias (i.e., weakened or strengthened speaker commitment to the modified proposition), see the discussion of 'modal spread' by Giannakidou & Mari (2018). Despite the numerous attempts to integrate MC into semantic theories, the empirical picture remains unclear: Do MC constructions really share the same semantics as their SM counterparts?

Another question that arises naturally is the use of MC in comparison to SM: Assuming MC and SM are 'functional equivalents', that is, variants of a linguistic variable with the same semantic interpretation (Labov 2006: 145-146), why would a speaker choose to use MC over SM? This question is specific to MC in the modal domain, as it is available across languages and varieties of English, in contrast to the phenomenon of 'multiple modals' (e.g., *might could*), which is much more restricted, that is, only possible in some varieties of English (see the first sentence in Lyons' 1977 quote above, and Kortmann & Schneider 2004 on colloquial American and Appalachian American English) and in general, in a smaller set of natural languages. Based on the sociolinguistic literature and more recent work on register in Lüdeling et al. (2022) and Pescuma et al. (2023), we assume that the choice between MC and SM can be register-driven, that is, the speaker chooses one or the other depending on the properties of the situational contexts with regard to, for example, interlocutor relations or communicative purposes. Furthermore, the choice can be better understood in relation to the social meaning, that is, "the set of inferences that can be drawn on the basis of how language is used in a specific interaction" (Hall-Lew et al. 2021; 3). For example, Glass (2015) reports on a social meaning study of three different semi-modals of universal force in English, see (2), and argues that "*you need to* has a slightly different semantics than *you have/got to*, and that this subtlety gives it a unique social meaning that depends on whether the speaker is licensed to tell the hearer what's good for him" (p.87). Similarly to that study, our study also addresses the social meaning of modal expressions of the same—existential or universal—force: More specifically, when a speaker utters MC in a specific context, what pragmatic inferences can be drawn about the speaker's backgrounds and personae?

(2) You {**need to /have to/gotta**} admire her for persevering. (Glass 2015: example (1))

In the following, we report on an experimental study on MC in English, which is guided by the following questions: (RQ1) How are MC constructions interpreted? We tested MC and SM constructions, using speaker commitment ratings, and found that their interpretations differ. (RQ2) How are MC constructions perceived? We tested the social meaning of MC and SM, as well as that of possibility and necessity modals, and found that their social meanings differ. (RQ3) How does the perception of doubling constructions differ? We will discuss the results of MC in the current study with the results of negative concord (NC) in our previous work (Rotter & Liu 2025,

2024) and show that the social meanings of MC and NC differ.

The paper is structured as follows: Section 2 presents the method and the results of the experiment. Section 3 discusses the results, the limitations and future outlooks of the study, and concludes the paper.

## 2. Experiment.

2.1. MATERIAL AND DESIGN. The experiment was implemented with PCIbex farm and hosted on its platform (https://farm.pcibex.net/). We used a 2x2-factorial design with the factors FORCE (possibility vs. necessity) and NUMBER (of used modal elements, i.e., MC vs. SM) in a Latin Square design with four lists. We used 12 critical, see (3), and 29 filler items of similar structure. (S1) remained the same across all items of the study, see (3). (S2) contained the FORCE and NUMBER manipulation. (Q1) was used to measure the interpretation of the sentence in terms of speaker commitment ratings (Liu et al. 2021) and (Q2) to measure the grammaticality ratings.

(3) (S1) Somebody says:
(S2) "I / **may possibly**$_{MC}$ / have gotten / the wrong address." (possibility)
"I / **may**$_{SM}$ / have gotten / the wrong address." (possibility)
"I / **must certainly**$_{MC}$ / have gotten / the wrong address." (necessity)
"I / **must**$_{SM}$ / have gotten / the wrong address." (necessity)
(Q1) Does the person believe they have gotten the wrong address?
(Q2) Is the sentence grammatical?

2.2. PROCEDURE. The experiment consisted of four parts: (**P1**) consent, (**P2**) two practice sentences to get familiarized with the experimental set-up, (**P3**) main study, and (**P4**) demographic and language background survey.

Participants were instructed to use the space bar to continue. Each trial in (**P2/3**) started with a fixation cross in the middle of the screen. Then (S1) appeared on the screen. (S2) was presented as a self-paced reading task; the length and position of each chunk was indicated by a line on the screen. Previous chunks disappeared so that only the current chunk was visible. Finally, the entire sentence (S2) appeared on the screen together with the following social meaning measures: **social backgrounds** of the speaker (*low*/*high socioeconomic status* and *low*/*high education*), and **persona** of the speaker (*in*/*formal*, *un*/*confident*, *im*/*polite*, *un*/*friendly*, *cold*/*warm*, *un*/*cool*, and *obedient*/*rebellious*); the measures appeared in groups of three and within each group, the order was randomized. In the two last screens, the participants first rated the speaker commitment, that is, **interpretation** (Q1-*Certainly no*/*Certainly* yes), and then the **grammaticality** of the sentence (Q2-*Certainly no*/*Certainly yes*). We chose this order to allow for unbiased speaker commitment ratings. All the measures and questions used a 7-point Likert scale in which the end- and midpoints were labels (e.g., 1: *ungrammatical* – 4: *undecided* – 7: *grammatical*). Participants used the mouse click to provide their answer.

2.3. PARTICIPANTS. We collected data from 104 native speaker of US English on the crowdsourcing platform Prolific (https://www.prolific.co/). The experiment took roughly 35 minutes and participants received monetary compensation. All participants provided their informed consent as approved by the Ethics Committee of the Deutsche Gesellschaft für Sprach-

wissenschaft (DGfS) in the context of SFB 1412 'Register'. Based on the inclusion criteria (see 2.4 for details), we analyzed the data of 93 participants from 37 of the 50 US states (mean age=39.2/SD=10.9, range=[19, 63]; female N=48, male N=43, non-binary N=2).

2.4. DATA ANALYSIS. The data was processed and analyzed with the open source software 'R' (version 4.1.1, R Core Team 2024) in the RStudio environment (RStudio Team 2024). We defined demographic (native English speaker and aged between 18 to 65 years) and attendance (75% of the chunks were attended for at least 150msec) inclusion criteria: First, all participants matched the demographic criteria. Second, the data of 11 participants were excluded based on the attendance. Furthermore, we removed the data of 254 (10.2%) trails, in which at least one chunk was attended for less than 150msec.

We calculated separate models for each measure with the remaining data in the cumulative link function model framework (Howcroft & Rieser 2021, Liddell & Kruschke 2018) using the package '`ordinal`' (Christensen 2019). The link function of each model was determined as the one with the highest log-likelihood value among the five possible link functions (i.e., probit, logit, cauchit, loglog, and cloglog) (Christensen 2019). Each model contained the sum-coded factors FORCE (possibility: $0.5$, necessity: $-0.5$) and NUMBER (MC: 0.5, SM: $-0.5$), as well as their two-way interaction FORCE×NUMBER (2-int). If the interaction turned out significant, we conducted a sub-analysis by splitting the data along the significant main factor. The sub-analysis included the same coding of the main effects.

We used the most parsimonious model approach to obtain the random effect structures; the used model structures are indicated in the respective result section. Log-likelihood ratio test comparisons of nested models (Bates et al. 2018) were used to obtain p-values, which were defined as significant at a value below 0.05. We reduced the complexity of models if their NULL-model did not converge. We used the model with the smaller Akaike Information Criterion (AIC) if the model choice was between subject or item intercepts. All statistical values of means, estimates and the like are rounded to the second decimals except for p-values smaller than 0.01.

**3. Results.** Below, we first present the results of the interpretation and grammaticality measures and then the results of the social meaning measures.

3.1. INTERPRETATION AND GRAMMATICALITY. The output of the models using interpretation (Q1) and grammaticality (Q2) ratings as dependent variable are shown in Table 2.

For the **interpretation** ratings, the logit-link function model including random subject intercepts and slopes for FORCE, NUMBER, and their interaction fit the data best. The results showed a significant main effect of FORCE in that possibility conditions were rated lower than necessity conditions ($\hat{\beta}$=$-$3.65, $\chi^2$(1)=95.99, p<0.001). There was a main effect of NUMBER in that MC was rated higher than SM ($\hat{\beta}$=0.60, $\chi^2$(1)=15.91, p<0.001). Furthermore, the 2-way interaction turned out significant ($\hat{\beta}$=$-$1.85, $\chi^2$(1)=41.51, p=<0.001), which results from a cross-over effect:

- Weakening effect in possibility modals: MC received significantly lower speaker commitment ratings than SM ($\hat{\beta}$=$-$0.35, $\chi^2$(1)=6.79, p<0.009)
- Strengthening effect in necessity modals: MC received significantly higher speaker commitment ratings than SM ($\hat{\beta}$=1.50, $\chi^2$(1)=36.92, p<0.001).

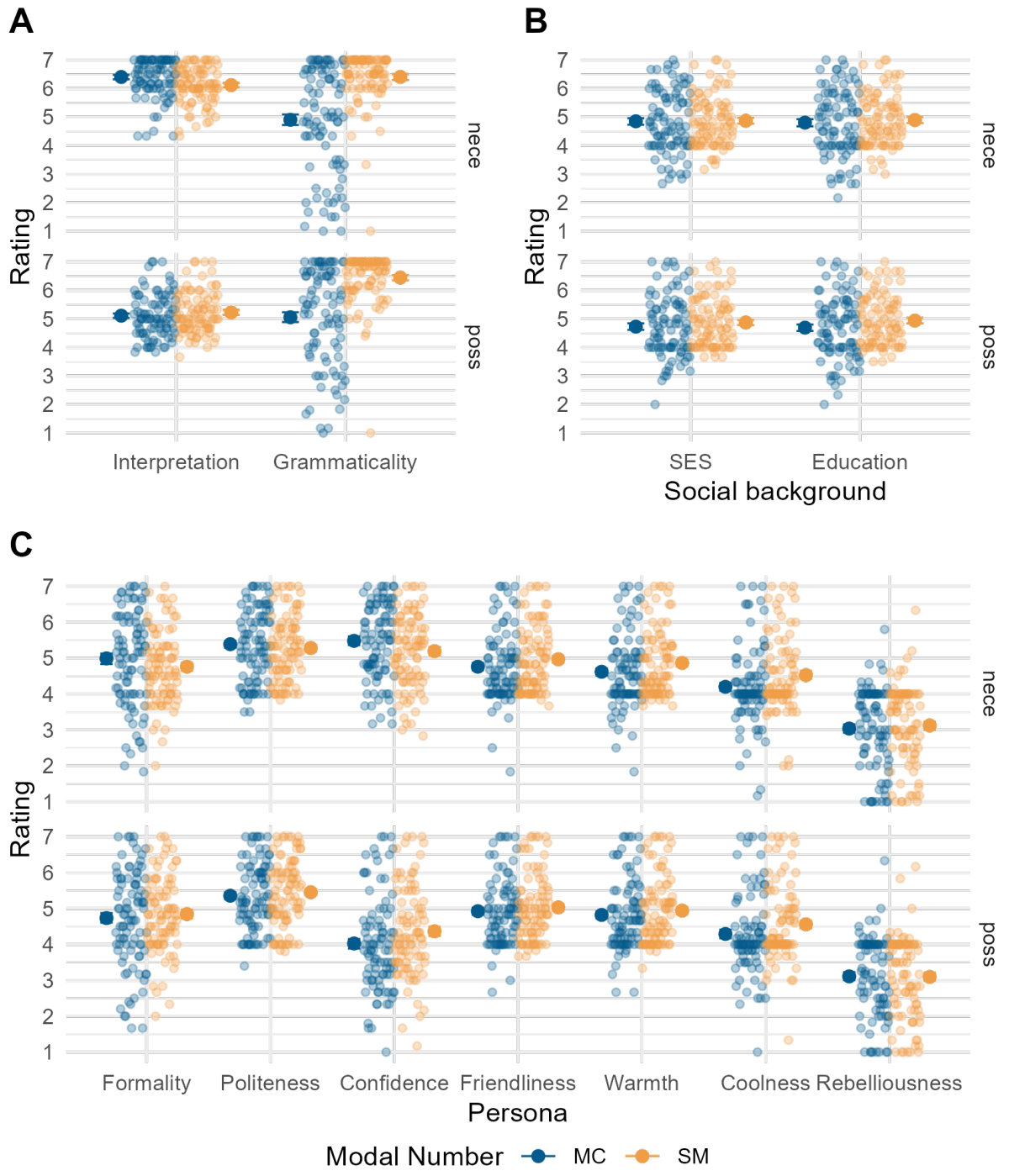


Figure 1: By-subject means (transparent dots) of the rating measures in comparison to overall means (opaque dots with error bars). **Panel A** depicts the ratings of the interpretation and grammaticality, **Panel B** those of the social background measures, and **Panel C** those of the the persona measures. The x-axis indicates the specific measures. The top scale shows the ratings from the *necessity* conditions (abbreviated as nece), the one below those of the *possibility* conditions (abbreviated as poss). The colors indicate the factor NUMBER with MC (i.e., *modal concord*) in blue and SM (i.e., *single modal*) in yellow. The y-axis depicts the ratings on a 7-point Likert scale. SES abbreviates *socioeconomic status*.

| | | n | Med | Mean | SD | SE | Med | Mean | SD | SE |
|---|---|---|---|---|---|---|---|---|---|---|
| FORCE | NUMBER | | **Interpretation** | | | | **Grammaticality** | | | |
| poss | MC | 542 | 5 | 5.11 | 0.96 | 0.04 | 6 | 5.06 | 2.05 | 0.09 |
| poss | SM | 541 | 5 | 5.22 | 0.98 | 0.04 | 7 | 6.45 | 1.09 | 0.05 |
| nece | MC | 536 | 7 | 6.40 | 0.90 | 0.04 | 6 | 4.90 | 2.18 | 0.09 |
| nece | SM | 539 | 6 | 6.12 | 0.85 | 0.04 | 7 | 6.41 | 1.16 | 0.05 |
| | | | **Socioeconomic status** | | | | **Education** | | | |
| poss | MC | 542 | 5 | 4.73 | 1.30 | 0.06 | 5 | 4.69 | 1.37 | 0.06 |
| poss | SM | 541 | 5 | 4.87 | 1.11 | 0.05 | 5 | 4.94 | 1.12 | 0.05 |
| nece | MC | 536 | 5 | 4.85 | 1.40 | 0.06 | 5 | 4.80 | 1.46 | 0.06 |
| nece | SM | 539 | 5 | 4.87 | 1.18 | 0.05 | 5 | 4.89 | 1.19 | 0.05 |
| | | | **Formality** | | | | | | | |
| poss | MC | 542 | 5 | 4.73 | 1.61 | 0.07 | | | | |
| poss | SM | 541 | 5 | 4.84 | 1.38 | 0.06 | | | | |
| nece | MC | 536 | 5 | 4.99 | 1.66 | 0.07 | | | | |
| nece | SM | 539 | 5 | 4.76 | 1.40 | 0.06 | | | | |
| | | | **Politeness** | | | | **Confidence** | | | |
| poss | MC | 542 | 5 | 5.36 | 1.17 | 0.05 | 4 | 4.03 | 1.61 | 0.07 |
| poss | SM | 541 | 5 | 5.45 | 1.11 | 0.05 | 4 | 4.36 | 1.51 | 0.06 |
| nece | MC | 536 | 5 | 5.39 | 1.16 | 0.05 | 6 | 5.48 | 1.41 | 0.06 |
| nece | SM | 539 | 5 | 5.28 | 1.15 | 0.05 | 5 | 5.19 | 1.39 | 0.06 |
| | | | **Friendliness** | | | | **Warmth** | | | |
| poss | MC | 542 | 5 | 4.92 | 1.17 | 0.05 | 5 | 4.82 | 1.19 | 0.05 |
| poss | SM | 541 | 5 | 5.03 | 1.14 | 0.05 | 5 | 4.94 | 1.16 | 0.05 |
| nece | MC | 536 | 5 | 4.76 | 1.21 | 0.05 | 4 | 4.62 | 1.25 | 0.05 |
| nece | SM | 539 | 5 | 4.97 | 1.15 | 0.05 | 5 | 4.86 | 1.17 | 0.05 |
| | | | **Coolness** | | | | **Rebelliousness** | | | |
| poss | MC | 542 | 4 | 4.29 | 1.33 | 0.06 | 3 | 3.11 | 1.33 | 0.06 |
| poss | SM | 541 | 4 | 4.56 | 1.27 | 0.05 | 3 | 3.09 | 1.34 | 0.06 |
| nece | MC | 536 | 4 | 4.20 | 1.32 | 0.06 | 3 | 3.04 | 1.34 | 0.06 |
| nece | SM | 539 | 4 | 4.53 | 1.27 | 0.05 | 3 | 3.12 | 1.36 | 0.06 |

Table 1: Descriptive statistics of the interpretation (Q1), grammaticality (Q2), social background, and persona ratings. Poss stands for *possibility* conditions, nece for *necessity*, MC for *modal concord*, and SM *single modal*. Med abbreviates *median*, SD *standard deviation*, and SE *standard error*.

For the **grammaticality** ratings, the logit-link function model including random subject and item intercepts, both with slopes for FORCE and NUMBER fit the data best. The results showed a significant main effect of NUMBER in that MC was rated less grammatical than SM ($\hat{\beta}$=−3.48, $\chi^2(1)$=83.01, p<0.001).

| | Fixed effects | | Model comparison | | Fixed effects | | Model comparison | |
|---|---|---|---|---|---|---|---|---|
| | $\hat{\beta}$ | SE | $\chi^2(1)$ | p-value | $\hat{\beta}$ | SE | $\chi^2(1)$ | p-value |
| | **Interpretation** | | | | **Grammaticality** | | | |
| FORCE | −3.65 | 0.30 | 95.99 | <0.001⋆ | 0.21 | 0.29 | 0.53 | 0.47 |
| NUMBER | 0.60 | 0.15 | 15.91 | <0.001⋆ | −3.48 | 0.36 | 83.01 | <0.001⋆ |
| 2-int | −1.85 | 0.28 | 41.51 | <0.001⋆ | 0.15 | 0.23 | 0.43 | 0.51 |

Table 2: Output of the model using interpretation and grammaticality ratings. 2-int abbreviates the *two way interaction FORCE×NUMBER*. SE stands for the *standard error*.

3.2. SOCIAL BACKGROUND. The output of the models using social background ratings as dependent variables are shown in Table 3.

| | Fixed effects | | Model comparison | | Fixed effects | | Model comparison | |
|---|---|---|---|---|---|---|---|---|
| | $\hat{\beta}$ | SE | $\chi^2(1)$ | p-value | $\hat{\beta}$ | SE | $\chi^2(1)$ | p-value |
| | **Socioeconomic status** | | | | **Education level** | | | |
| FORCE | −0.15 | 0.05 | 9.13 | <0.003⋆ | −0.11 | 0.05 | 4.84 | < 0.03⋆ |
| NUMBER | −0.01 | 0.05 | 0.01 | 0.92 | −0.06 | 0.05 | 1.25 | 0.26 |
| 2-int | −0.26 | 0.10 | 6.58 | 0.01⋆ | −0.31 | 0.10 | 9.62 | <0.002⋆ |

Table 3: Output of the model using the social background ratings. 2-int abbreviates the *two way interaction FORCE×NUMBER*. SE stands for the *standard error*.

For the **SES** and **education level** ratings, the loglog-link function model including random subject intercepts fit the data best. The results showed a significant main effect of FORCE in that possibility conditions received lower SES ($\hat{\beta}$=−0.15, $\chi^2(1)$=9.13, p<0.003) and education level ($\hat{\beta}$=−0.11, $\chi^2(1)$=4.84, p<0.03) ratings than necessity conditions. Furthermore, the 2-way interaction showed a significant effect in the SES model (SES-model: $\hat{\beta}$=−0.26, $\chi^2(1)$=6.58, p=0.01; education level-model: $\hat{\beta}$=−0.31, $\chi^2(1)$=9.62, p<0.002), which results from MC receiving lower SES ($\hat{\beta}$=−0.18, $\chi^2(1)$=5.55, p<0.02) and education level ($\hat{\beta}$=−0.26, $\chi^2(1)$=11.92, p<0.001) ratings than SM only in possibility conditions.

3.3. PERSONAE. The output of the models using personae ratings as dependent variables are shown in Table 4.

For the **formality** ratings, the cloglog-link function model including random subject intercepts with FORCE, NUMBER, and their 2-way interaction as well as random item intercepts fit the data best. The results showed a significant main effect of FORCE in that possibility conditions received lower formality ratings than necessity conditions ($\hat{\beta}$=−0.18, $\chi^2(1)$=4.19, p=0.04). Furthermore, the 2-way interaction showed a significant effect ($\hat{\beta}$=−0.50, $\chi^2(1)$=9.12, p<0.003), which results from MC receiving higher formality ratings than SM only in necessity conditions ($\hat{\beta}$=0.45, $\chi^2(1)$=8.80, p=0.003).

For the **politeness** ratings, the cloglog-link function model including random subject intercepts fit the data best. The results showed no significant effects.

For the **confidence** ratings, the logit-link function model including random subject and item intercepts fit the data best. The results showed a significant main effect of FORCE in that possibil-

| | Fixed effects | | Model comparison | | Fixed effects | | Model comparison | |
|---|---|---|---|---|---|---|---|---|
| | $\hat{\beta}$ | SE | $\chi^2(1)$ | p-value | $\hat{\beta}$ | SE | $\chi^2(1)$ | p-value |
| | **Formality** | | | | | | | |
| FORCE | −0.18 | 0.09 | 4.19 | 0.04⋆ | | | | |
| NUMBER | 0.21 | 0.12 | 3.17 | 0.08 | | | | |
| 2-int | −0.50 | 0.16 | 9.12 | <0.003⋆ | | | | |
| | **Politeness** | | | | **Confidence** | | | |
| FORCE | 0.08 | 0.06 | 2.06 | 0.15 | −1.79 | 0.09 | 453.50 | <0.001⋆ |
| NUMBER | 0.03 | 0.06 | 0.22 | 0.64 | 0.05 | 0.08 | 0.45 | 0.50 |
| 2-int | −0.21 | 0.11 | 3.56 | 0.06 | −1.03 | 0.16 | 41.28 | <0.001⋆ |
| | **Friendliness** | | | | **Warmth** | | | |
| FORCE | 0.14 | 0.05 | 7.16 | <0.008⋆ | 0.16 | 0.05 | 9.47 | 0.002⋆ |
| NUMBER | −0.20 | 0.05 | 14.01 | <0.001⋆ | −0.26 | 0.05 | 23.74 | <0.001⋆ |
| 2-int | 0.08 | 0.10 | 0.52 | 0.47 | 0.13 | 0.10 | 1.62 | 0.20 |
| | **Coolness** | | | | **Rebelliousness** | | | |
| FORCE | 0.02 | 0.05 | 0.15 | 0.70 | 0.02 | 0.05 | 0.10 | 0.76 |
| NUMBER | −0.42 | 0.05 | 62.42 | <0.001⋆ | −0.05 | 0.07 | 0.45 | 0.50 |
| 2-int | 0.20 | 0.10 | 3.62 | 0.06 | 0.13 | 0.10 | 1.91 | 0.17 |

Table 4: Output of the model using the social background ratings. 2-int abbreviates the *two way interaction FORCE×NUMBER*. SE stands for the *standard error*.

ity conditions received lower ratings than necessity conditions ($\hat{\beta}$=−1.79, $\chi^2$(1)=453.50, p<0.001). Furthermore, the 2-way interaction turned out significant ($\hat{\beta}$=−1.03, $\chi^2$(1)=41.28, p<0.001), which results from a cross-over effect: MC received significantly lower confidence ratings in possibility conditions than SM ($\hat{\beta}$=−0.66, $\chi^2$(1)=20.25, p<0.001), while MC received higher ratings than SM in necessity conditions ($\hat{\beta}$=0.52, $\chi^2$(1)=16.76, p<0.001).

For the **friendliness** and **warmth** ratings, the loglog-link model with random subject intercepts fit the data best. The results showed a significant main effect of FORCE in that possibility conditions received higher friendliness ($\hat{\beta}$=0.14, $\chi^2$(1)=7.16, p<0.008) and warmth ($\hat{\beta}$=0.16, $\chi^2$(1)=9.47, p=0.002) ratings than necessity conditions. There was a significant main effect of NUMBER in that MC received lower friendliness ($\hat{\beta}$=−0.20, $\chi^2$(1)=14.01, p<0.001) and warmth ($\hat{\beta}$=−0.26, $\chi^2$(1)=23.74, p<0.001) ratings than SM conditions.

For the **coolness** ratings, the loglog-link model with random subject intercepts fit the data best. The results showed a significant effect of NUMBER in that MC received lower coolness ratings than SM ($\hat{\beta}$=−0.42, $\chi^2$(1)=62.42, p<0.001).

For the **rebelliousness** ratings, the probit-link model with random subject intercepts and slopes for NUMBER fit the data best. The results showed no significant effects.

**4. Discussion and Conclusion.** The results of the study are summarized in Table 5.

First, MC was rated as less grammatical (with the mean ratings 5.06/4.90 for possibility/necessity MC) than SM. We take this to be in line with the more restricted distribution of MC.

In relation to (RQ1), our study shows that MC conveys a weakened speaker commitment

| | FORCE | NUMBER | 2-int FORCE×NUMBER | |
|---|---|---|---|---|
| **Interpretation** | poss < ness | MC > SM | poss: MC < SM | ness: MC > SM |
| **Grammaticality** | – | MC < SM | – | – |
| **Socioeconomic status** | poss < ness | – | poss: MC < SM | – |
| **Education level** | poss < ness | – | poss: MC < SM | – |
| **Formality** | poss < ness | – | – | ness: MC > SM |
| **Politeness** | – | – | – | – |
| **Confidence** | poss < ness | – | poss: MC < SM | ness: MC > SM |
| **Friendliness** | poss > ness | MC < SM | – | – |
| **Warmth** | poss > ness | MC < SM | – | – |
| **Coolness** | – | MC < SM | – | – |
| **Rebelliousness** | – | – | – | – |

Table 5: Summary of the results. The abbreviation 'poss' stands for *possibility* and 'ness' for *necessity* conditions. MC stands for *modal concord* and SM for *single modal*. 2-int abbreviates the *two-way interaction* FORCE×NUMBER. The symbol < indicates that the left entity is smaller than the right one; > represents the reverse. The symbol – indicates the lack of a significant effect.

than SM in possibility modals and a strengthened speaker commitment in necessity modals. This finding challenges the semantic equivalence assumption for MC and SM as well as the claim that modal adverbs in MC constructions have no semantic contributions. One way out would be to analyze weakening or strengthening as a pragmatic effect of, say, "reinforcing" (Halliday 1970, see the quote above from Lyons 1977) of the respective modal force (i.e., existential or universal): in simple words, doubling does not add extra contributions to the semantics but give rise to enriched meanings about speaker assumptions. Alternatively, such effects can be taken as arguments for the 'modal spread' analysis of Giannakidou & Mari (2018), by which the weakening and strengthening effect can be analyzed as semantic contributions of modal adverbs. While we do not provide any formal analysis for this in the paper, we see several arguments in favor of the 'modal spread' analysis. Firstly, if we take modal adverbs such as *certainly* in *must certainly* to have no semantics in MC constructions, we need to stipulate a different lexical entry for the same adverb in other contexts, for example, where it co-occurs with a modal verb of a different force. Our preliminary corpus study shows that (4-a) is a naturally-occurring combination, which is "modally non-harmonic" (see the quote above from Lyons 1977)—in contrast, the well-formedness of (4-b) is questionable, a topic we leave for future work. Given the well-formedness of (4-a) as well as cases where modal adverbs occur without verbs, it strikes us as unattractive to analyze modal adverbs as semantically vacuous for MC constructions.

(4) a. It **may certainly** be the case.
b. ?/*It **must possibly** the base.

In relation to (RQ2), (i) MC was rated as less friendly, warm, and cool than SM. This is something which we neither predicted nor have a good explanation for, and thus needs to be further investigated in future studies. Crucially, certain measures showed a significant NUMBER*FORCE interaction in that (ii) in possibility modals, MC was rated as significantly lower than SM in SES,

education and confidence levels and (iii) in necessity modals, MC was rated as more formal and more confident than SM. Relating these results with the results of the interpretation, it seems that weaker statements give rise to more negative perceptions and stronger ones to more positive perceptions, indicating a positive correlation between meaning strength and social perception.

In relation to (RQ3), we only compare MC with NC here, see a brief grammatical comparison of these two doubling phenomena in Zeijstra (2007). Sociolinguistics and recent research on social meaning have both documented work on NC, which is more prevalent in different varieties of English and socially stigmatized (Eckert 1989, 2019, Moore 2021, Rotter & Liu 2025). In comparison, there is little work on the social meaning of MC, which unlike NC, has a wide distribution in English(es) and is not socially stigmatized. In recent work, Rotter & Liu (2025, 2024) conducted a series of studies on the register-sensitivity and the social meaning of NC compared to negative polarity items (NPIs) in American English, see (5). The findings (in the data by mostly self-reported non-dialect speakers of US English) are that (i) NC use is register-sensitive in that it was rated less appropriate in formal than in informal contexts, (ii) comprehenders interpreted NC similarly to NPIs, but (iii) associated NC use with lower socioeconomic and education levels, as well as less formal but more rebellious personae of the speaker. The current study, in combination with our previous work, provides experimental evidence that the doubling phenomena of MC and NC have very distinct social meanings.

(5) a. I didn't see **nothing**. (NC construction)
b. I didn't see **anything**. (NPI construction)

Before we finish the paper, we would like to briefly address the limitations and future outlooks of our study. First, our study compared MC with SM constructions with a single modal auxiliary. The design can be extended to the comparison between MC, SM-verb and SM-adverb (e.g., *must certainly* vs. *must* vs. *certainly* and *may possibly* vs. *may* vs. *possibly*). Secondly, we used isolated sentences in the study, which was our first step towards understanding the social meaning of MC, but it is evident that the set of inferences to be drawn from its use can differ greatly depending on the specific interaction. The next step would therefore be to incorporate context and further explore the question of whether MC is register-sensitive, and what social meanings arise in relation to its use in different situational contexts.

**Data repository.** The data set for this study can be found in the online repository: `https://osf.io/47rkg/`.

**Authorship contribution statement. Mingya Liu:** Conceptualization, Formal analysis, Methodology, Writing - original draft, Supervision, Project administration, Funding acquisition. **Stephanie Rotter:** Methodology, Data collection, Data analysis and visualization, Writing - original draft.

## References

Bates, Douglas, Reinhold Kliegl, Shravan Vasishth & Harald Baayen. 2018. Parsimonious mixed models. *arXiv* 1506. https://doi.org/10.48550/arXiv.1506.04967.

Christensen, Rune Haubo Bojesen. 2019. Ordinal: Regression models for ordinal data `CRAN.R-project.org/package=ordinal`. R package version 2023.12-4.

Eckert, Penelope. 1989. *Jocks and burnouts: Social categories and identity in the high school*. New York: Teachers College Press.

Eckert, Penelope. 2019. The limits of meaning: Social indexicality, variation, and the cline of interiority. *Language* 95(4). 751–776. https://doi.org/10.1353/lan.0.0239.

Geurts, Bart & Janneke Huitink. 2006. Modal concord. In Paul Dekker & Hedde Zeijlstra (eds.), *Proceedings of the ESSLLI 2006 workshop concord phenomena at the syntax-semantics interface*, 15–20. Malaga: University of Malaga.

Giannakidou, Anastasia & Alda Mari. 2018. The semantic roots of positive polarity: Epistemic modal verbs and adverbs in English, Greek and Italian. *Linguistics and Philosophy* 41(6). 623–664. https://doi.org/10.1007/s10988-018-9235-1.

Glass, Lelia. 2015. Strong necessity modals: Four socio-pragmatic corpus studies. In Paul Dekker & Hedde Zeijlstra (eds.), *University of Pennsylvania working papers in linguistics*, vol. 21(2), 79–88.

Hall-Lew, Lauren, Emma Moore & Robert J. Podesva. 2021. Social meaning and linguistic variation: Theoretical foundations. In Lauren Hall-Lew, Emma Moore & Robert J. Podesva (eds.), *Social meaning and linguistic variation*, 1–24. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108578684.001.

Halliday, Michael A. K. 1970. Functional diversity in language as seen from a consideration of mood and modality in English. *Foundations of Language* 6. 322–361.

Howcroft, David M. & Verena Rieser. 2021. What happens if you treat ordinal ratings as interval data? Human evaluations in NLP are even more under-powered than you think. In Marie-Francine Moens, Xuanjing Huang, Lucia Specia & Scott Wen-tau Yih (eds.), *Proceedings of the 2021 conference on empirical methods in natural language processing*, 8932–8939. Online and Punta Cana, Dominican Republic: Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.703. https://aclanthology.org/2021.emnlp-main.703. Accessed 2022-01-03.

Kortmann, Bernd & Edgar W. Schneider. 2004. *A multimedia reference tool. volume 1: Phonology. volume 2: Morphology and syntax*. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110197181.

Labov, William. 2006. *The Social Stratification of English in New York City*. Cambridge: Cambridge University Press 2nd edn. https://doi.org/10.1017/CBO9780511618208.

Liddell, Torrin M. & John K. Kruschke. 2018. Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology* 79. 328–348. https://doi.org/10.1016/j.jesp.2018.08.009.

Liu, Mingya, Stephanie Rotter & Anasatasia Giannakidou. 2021. Bias and modality in conditionals: Experimental evidence and theoretical implications. *Journal of Psycholinguistic Research* 50. 1369–1399. https://doi.org/10.1007/s10936-021-09813-z.

Lyons, John. 1977. *Semantics*, vol. 2. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511620614.

Lüdeling, Anke, Artemis Alexiadou, Aria Adli, Karin Donhauser, Malte Dreyer, Markus Egg, Anna Helene Feulner, Natalia Gagarina, Wolfgang Hock, Stefanie Jannedy, Frank Kammerzell, Pia Knoeferle, Thomas Krause, Manfred Krifka, Silvia Kutscher, Beate Lütke,

Thomas McFadden, Roland Meyer, Christine Mooshammer, Stefan Müller, Katja Maquate, Muriel Norde, Uli Sauerland, Stephanie Solt, Luka Szucsich, Elisabeth Verhoeven, Richard Waltereit, Anne Wolfsgruber & Lars Erik Zeige. 2022. Register: Language Users' Knowledge of Situational- Functional Variation– Frame Text of the First Phase Proposal for the CRC 1412. *Register Aspects of Language in Situation (REALIS)* 1(1). 1–57. https://doi.org/10.18452/24901.

Moore, Emma. 2021. The social meaning of syntax. In Lauren Hall-Lew, Emma Moore & Robert J.Editors Podesva (eds.), *Social meaning and linguistic variation: Theorizing the third wave*, 54–79. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108578684.003.

Pescuma, Valentina N., Diana Serova, Julia Lukassek, Antje Sauermann, Roland Schäfer, Aria Adli, Felix Bildhauer, Markus Egg, Kristina Hülk, Aine Ito, Stefanie Jannedy, Valia Kordoni, Milena Kühnast, Silvia Kutscher, Robert Lange, Nico Lehmann, Mingya Liu, Beate Lütke, Katja Maquate, Christine Mooshammer, Vahid Mortezapour, Stefan Müller, Muriel Norde, Elizabeth Pankratz, Angela G. Patarroyo, Ana-Maria Plesca, Camilo Rodríguez-Ronderos, Stephanie Rotter, Uli Sauerland, Gohar Schnelle, Britta Schulte, Gediminas Schüppenhauer, Bianca Maria Sell, Stephanie Solt, Megumi Terada, Dimitra Tsiapou, Elisabeth Verhoeven, Melanie Weirich, Heike Wiese, Kathy Zaruba, Lars Erik Zeige, Anke Lüdeling & Pia Knoeferle. 2023. Situating language register across the ages, languages, modalities, and cultural aspects: Evidence from complementary methods. *Frontiers in Psychology* 13. https://doi.org/10.3389/fpsyg.2022.964658.

R Core Team. 2024. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. `https://www.R-project.org/`.

Rotter, Stephanie & Mingya Liu. 2024. A register approach to negative concord vs. negative polarity items in English. *Linguistics* https://doi.org/10.1515/ling-2023-0016.

Rotter, Stephanie & Mingya Liu. 2025. Less formal and more rebellious — an experiment on the social meaning of negative concord in American English. In Chiara Gianollo & Johan van der Auwera (eds.), *Negative concord: A hundred years on*, vol. 385, 303–332. Berlin, Bosten: De Gruyter Mouton.

RStudio Team. 2024. *Rstudio: Integrated development environment for r*. RStudio, PBC Boston, MA. `http://www.rstudio.com/`.

Yalcin, Seth. 2007. Epistemic modals. *Mind* 116(464). 983–1026. https://doi.org/10.1093/mind/fzm983.

Zeijstra, Hedde. 2007. Modal concord. *Proceedings of Semantics and Linguistic Theory* (17). 317–332. https://doi.org/10.3765/salt.v17i0.2961.