Definitely, maybe: Approaching speaker commitment experimentally

When speakers are certain about some fact expressed by a proposition *p*, they are likely to communicate this fact with a simple declarative utterance, as in (1a). When they want to convey epistemic uncertainty, they have different devices for doing so: e.g., using a modal adverbial (1b) or a modal verb (1c). Either option leaves open the possibility that *p* is not true.

(1) a. The coffee is cold. b. The coffee is probably cold.
c. The coffee might be cold. d. The coffee must be cold.

While it is uncontroversial that epistemic uncertainty in both (1b) and (1c) is conveyed by clearly lexical means, it is still debated whether the uncertainty component in (1d) reduces to the same interpretive mechanism. The relative weakness of *must* in (1d), compared to the simple declarative in (1a), is puzzling because *must* serves as a strong modal of necessity (cf. Karttunen 1972). How do listeners understand that *p* does not follow from *must p*? One possible solution to this semantic puzzle is that the weakness of epistemic *must* need not derive from truth-conditional content (e.g., quantification over a limited set of possible worlds). Rather, the epistemic interpretation of *must* may rely on contextual factors like evidence strength and thus may involve additional (pragmatic) reasoning of the listener.

Languages besides English have many more ways to convey weakened speaker commitment. Even restricting ourselves to Germanic languages, we observe interesting differences. For instance in German, many devices for expressing speaker certainty do not differ from English: (2a-c) are the German equivalents of (1a, b, d). However, (2d) is different:

(2) a. Der Kaffee ist kalt geworden. b. Der Kaffee ist vermutlich kalt geworden.
c. Der Kaffee muss kalt geworden sein. d. Der Kaffee ist wohl kalt geworden.

Discourse particles like *wohl* in (2d) organize the discourse by conveying the epistemic states of both the speaker and the hearer. What is interesting about (2d) in our context is that discourse particles such as *wohl*, although lexically encoding weakened speaker commitment, are not part of the truth-conditional content of an utterance, in contrast to modal adverbs such as *vermutlich* ('presumably', 2b). Just as the epistemic use of *must*/German *muss* (1d)/(2c), *wohl* is thus a special case of expressing weakened speaker commitment. Specifically, Zimmermann (2004, 2008) has shown that this particular particle cannot affect the presuppositional domain of utterance interpretation, unlike modal adverbs such as *vermutlich*.

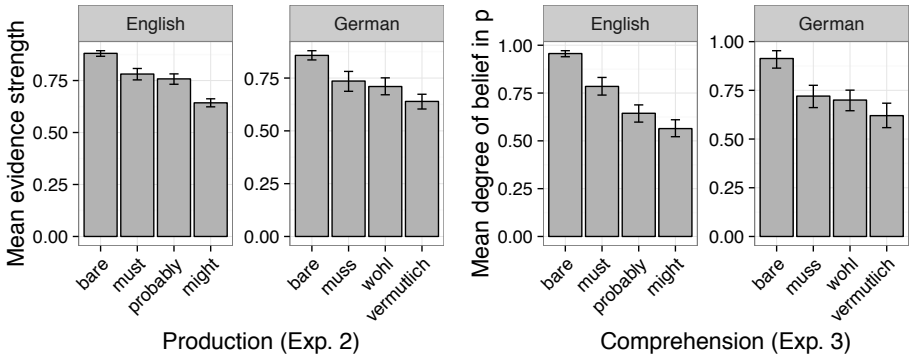
Given this cross-linguistic sketch of expressions conveying different strengths of speaker commitment, we addressed the following questions experimentally: (i) Under which evidential circumstances (measured in Exp. 1) do speakers prefer to use which evidential devices (Exp. 2)? (ii) Do listeners ascribe different strengths of speaker commitment to the use of these various linguistic devices (Exp. 3)?

In **Exp.1** (n=80; 40 English; 40 German), we collected estimates of evidence strength, which were used in the analysis of Exps. 2 and 3. Participants rated the probability of *p* (e.g., of 'It's raining' or German 'Es regnet') given a piece of evidence (e.g., *You hear the sound of water dripping on the roof*) on a sliding scale with endpoints labeled 'impossible' and 'certain', for five different pieces of evidence for four different *p* each (20 total). The two populations (English and German) did not differ in their estimates of evidence strength.

In **Exp.2** (n=80; 40 English; 40 German), we tested how likely speakers are to use one of the four possible utterances in (1) and (2) as evidence strength decreases. On each trial, participants were presented with a piece of evidence (e.g., *You see a person come in from outside with wet hair and wet clothes*) and were asked to choose one utterance to describe the situation to a friend. The main question of interest is whether the choice of form to communicate about *p* depends on the strength of the evidence for *p*. **RESULTS.** In two separate mixed-effects linear regressions (one for the English, one for the German data) predicting evidence strength from a dummy-coded predictor for utterance choice (with *must*/*muss* as

reference level), evidence strength was greater when bare *p* was produced than when *must*/*muss p* was produced in both English and German. In English, evidence was weaker when *might p* was produced. There was no difference in evidence strength between *must p* and *probably p*. In German, evidence was weaker when *vermutlich p* was produced. Interestingly, there was no difference in evidence strength between *muss p* and *wohl p*. Accordingly, speakers tend to use *muss* and *wohl* instead of the respective adverb when evidence strength is greater.

Exp.3 (n=120; 60 English; 60 German) tested whether listeners' estimates of speakers' belief in *p* and strength of evidence for *p* vary by observed utterance. On each trial, participants saw an utterance (e.g., *It's raining*), and were asked (i) to judge whether the speaker believes *p* on a sliding scale with endpoints labeled 'definitely not' (coded as 0) and 'definitely' (coded as 1); and (ii) to select one of five pieces of evidence the speaker must have had for *p*. **RESULTS.** Participants rated speakers as more likely to believe *p* after observing bare *p* than after observing *must p* in both English and German. In English they believed *p* was less likely after observing either *probably p* and *might p*. In German, there were no differences in ascribed degree of belief in *p* between *muss p* and *wohl p,* but belief in *p* was rated as less likely after *vermutlich p*. These results mirror the evidence strength effects found in production (Exp.2). Speakers tend to use epistemic *must*/*muss* instead of corresponding modals when the degree of evidence strength is greater; and listeners take this into account in interpretation: ascribed speaker commitment in the case of both epistemic *must* and discourse particles is stronger than in the case of using otherwise synonymous adverbs such as *vermutlich*.



Our results confirm well-known differences of expressing different strengths of speaker commitment (von Fintel & Gillies 2010; Lassiter 2016). The epistemic use of *must*/*muss* expresses a weaker commitment than the bare form, and the use of the discourse particle *wohl* conveys a stronger commitment than otherwise synonymous modal adverbs. However, we also observe unpredicted parallel effects of epistemic *must*/*muss* and the particle *wohl* (both in production and in comprehension). We suggest that these two devices pattern alike because both epistemic *must*/*muss* and *wohl* involve a different interpretive route that listeners must take in order to arrive at the relevant reading of weakened speaker commitment. In the case of *wohl*, the particle leaves the asserted state of affairs untouched, and we thus expect the speaker to be more committed to the truth of the proposition when using a particle than when using a synonymous adverb expressing weakened commitment. In the case of epistemic *must*, our experiments lend support to the idea that *must* crucially differs from other modal devices such as *might* and *probably*. The weakening effect may thus be a non-lexical/non-propositional component of utterance interpretation. Assuming that a) the bare form is the canonical way of expressing *p* when the speaker's certainty about *p* exceeds a mutually agreed-upon threshold and b) the cost of uttering *must p* is greater than the bare form, we suggest that listeners infer both that the probability of *p* is smaller than when the utterance is the less costly bare form and that the speaker has less than maximally strong evidence for *p*. Thus, the semantics of *must* remains strong and pragmatics does the job of weakening.