Experimental Evidence for a Semantic Account of Free Choice

1 Introduction

Unlike plain disjunctive statements, modalized disjunctions such as (1) often lead to the inference that each disjunct is true: from an utterance of (1), one may conclude both that Mary may have a burger and that she may have steak.

(1) Mary may have a burger or (she may have) a steak.

This phenomenon is called *Free Choice* (FC) (Von Wright, 1968; Kamp, 1973). Since such readings are not predicted by classical analyses of disjunction and modality, the first task for a theory of FC is to explain why they are possible at all. Furthermore, there seem to be restrictions on their availability, which also call for an explanation. First, FC readings are sometimes unavailable in contexts where the speaker has limited knowledge about each disjunct. For example, if (1) is followed by "I forgot which", it would not receive a FC interpretation (Zimmerman, 2000). We call this factor *Speaker Knowledge* and we define a knowledgeable speaker (KS) as a speaker who knows whether each individual disjunct is true or false. The second fact that a theory of FC must explain is that the availability of a FC reading seems to be dependent on the scope of disjunction relative to the modal. For example, a narrow scope (NS) disjunction as in (2a) yields a FC reading, while a wide scope (WS) disjunction as in (2b) has been argued not to (Fox, 2007).

(2) a. Mary may either have ice-cream or cake.
b. Either Mary may have ice-cream or she may have cake.

Analyses of FC come in two varieties. Semantic approaches analyze FC effects using novel denotations for modals or disjunction (Zimmerman, 2000), while pragmatic approaches treat FC interpretations as implicatures (Kratzer & Shimoyama, 2002). In this paper, we test the predictions of analyses from each camp.

Aloni (2016): This semantic account of FC draws on nonstandard logics (Hawke & Steinert-Threlkeld, 2016). Propositions are evaluated against a state s (non-empty set of worlds, representing an agent's knowledge) instead of a single world w. A disjunction $\phi \vee \psi$ is supported by s iff s is divisible into two potentially overlapping substates, each of which supports one of the disjuncts. A modal statement $\diamond\phi$ is supported at s iff ϕ is supported by each state created by taking the intersection of the informational content of ϕ and the worlds accessible from some $w \in s$.

This account improves on Zimmerman's in that it can derive FC for NS disjunction. It predicts that the availability of FC readings depends on an interaction between scope and the *indisputability* of the accessibility relation. For deontic modals, indisputability corresponds to the pretheoretic notion of KS discussed above. Concretely, only WS modalized disjunctions in NON-KS contexts would lack a FC reading.

Fox (2007): This approach derives FC as a Scalar Implicature (SI) using a contradiction-free version of the *exhaustivity operator* introduced by Chierchia, 2004 which applies recursively to the set of alternatives for disjunction suggested by Sauerland, 2004. Applying the *exh* operator to (2a) with the set of alternatives in (3b) only yields an implicature that Mary can't have both foods at the same time, together with ignorance as to which one she can actually have. This uncertainty reading may seem implausible if the hearer knows that the speaker is knowledgeable. She may then reparse the utterance to derive a FC reading instead, by applying *exh* recursively, yielding the new set of alternatives C' (as shown in (3c)). Unlike C, all alternatives in C' can be negated without contradiction, and provide the FC interpretation in (4).

(3) a. $\diamond(p \vee q)$
b. $C = \{\diamond p, \diamond q, \diamond(p \wedge q)\ \}$
c. $C' = \{exh_C[\diamond p], exh_C[\diamond q], exh_C[\diamond(p \wedge q)]\ \}$

$= \{(\diamond p \wedge \neg \diamond q), (\diamond q \wedge \neg \diamond p), \diamond(p \wedge q)\}$

(4) $\diamond(p \vee q) \wedge \neg(\diamond p \wedge \neg \diamond q) \wedge \neg(\diamond q \wedge \neg \diamond p) \wedge \neg \diamond (p \wedge q) \models (\diamond p \wedge \diamond q)$

In this pragmatic account, the *opinionatedness* of the speaker determines whether an uncertainty or a FC interpretation arises. This essentially amounts to our Speaker Knowledge factor. Interestingly, this analysis does not extend to wide-scope disjunctions, which always receive uncertainty readings. Fox, 2007 argues that this is a good prediction because explicit wide-scope disjunctions do not seem to give rise to FC readings.

1.1 Summary

The predictions of Aloni, 2016 and Fox, 2007 are summarized in Table 1. The goal of our experiments was to test these competing predictions.

	non-KS		KS	
	NS	WS	NS	WS
Aloni, 2016	✓	×	✓	✓
Fox, 2007	×	×	✓	×

Table 1: Availability of FC readings in each configuration according to each theory

2 Experiment 1

2.1 Design

Our first experiment took the form of a pragmatic acceptability judgment task. At the beginning of the experiment, participants were shown a background story about a character named Danny; experimental items consisted of a context and a sentence. Participants were asked "Could Danny say that?" and responded using a 5-point Likert scale from "Not at all" to "Definitely". The two factors of Table 1 were manipulated within-subject. The context determined whether the speaker was knowledgeable or not (**KS** vs. **nKS**) and the utterance could involve narrow-scope or wide-scope disjunction (**NS** vs. **WS**).

(5) a. **nKS:** Danny only knows that Mary is not a 3rd grader.
b. **KS1:** Danny knows exactly which grade Mary is in.
c. **KS2:** Mary is not a 3rd grader and Danny knows whether Mary is a 1st or a 2nd grader.

(6) a. **NS:** Mary can have either a pizza or a hamburger.
b. **WS:** Either Mary can have a pizza or she can have a hamburger.

Fox, 2007, and more recently in Meyer and Sauerland, 2016, discuss the possibility that WS be reinterpreted as NS via across-the-board movement (Larson, 1985). To enforce the desired scope, we used the word 'either', which is known to block such reinterpretation.

To manipulate whether the context was **KS** or **nKS**, we created a background story where Danny is a cook working at an elementary school which has strict rules about what students can have for lunch. First graders can only have a pizza, second graders can only have hamburgers, but third graders can choose between the two foods. Therefore, knowing what a child may have for lunch amounts to knowing which grade they are in. This allowed us to manipulate Danny's knowledge regarding modal statements simply by manipulating his knowledge of which grade children are in, as illustrated in (5).

In each context, the choice of a reading (FC or uncertainty) affected the truth value of the sentence, and thus its expected rating. In the **nKS** context, the target sentences are plain false under a FC reading, and true under a non-FC reading. In particular, the primary implicatures of the non-FC reading are satisfied. To test for FC reading, we thus simply had to compare the target sentences to unambiguously true control sentences. To keep the comparison as minimal as possible, we used plain disjunctive statements in context (7a). The word *either* could appear either low (8a) or high (8b), mimicking the NS/WS contrast (although there was nothing to scope over).

In **KS** contexts, the ignorance implicatures of the non-FC reading are always violated. This means that comparing a target sentence to a base-

line could only lead to a contrast between a pragmatic violation and either a clear true or a clear false case. The solution we adopted was to use two contexts. In (5b), the target sentences are true under a FC reading (one would need to assume that Mary is a 3rd grader) and true but underinformative otherwise (henceforth u.i.). In (5c), the target sentences are false under a FC reading, and u.i. otherwise. Therefore, under a FC reading, the target sentences should receive maximally different ratings (plain true in **KS1**, plain false in **KS2**), but without a FC reading both sentences should receive the same, intermediate rating (corresponding to a u.i. sentence). As an added control, and to counterbalance for the fact that the sentences (8) would always be true otherwise, we created **UI** controls by combining them with context (7b).

(7) a. **TRUE (=nKS):** Danny only knows that Mary is not a 1st grader.
b. **UI:** Danny knows that Mary is a 3rd grader.

(8) a. **NS:** Mary is either a 2nd grader or a 3rd grader.
b. **WS:** Either Mary is a 2nd grader or she is a 3rd grader.

In short, our design allowed us to have a diagnosis for FC readings in **KS** and **nKS** contexts, and in both cases FC was indicated by a maximal difference between two conditions, while the absence of a FC reading was indicated by equal (un)acceptability of the two conditions. This is illustrated in Table 2.

We also included true and false fillers, which did not involve any modal or disjunction, and four training items with feedback right after the instructions. Each target condition and disjunctive control was repeated eight times, and each filler four times, for a total of 88 items. Fifty participants were recruited on MTurk; five were removed because of high error rates, and one for not reporting English as their native language.

2.2 Results

The data is presented in Figure 1a. After normalization, responses were analyzed with linear mixed-effects models following the procedure suggested in Bates, Kliegl, Vasishth, and Baayen, 2015 regarding the random effects structure.

We first ran the planned analysis on the data from **nKS** targets, Disjunction true controls and the two types of KS targets. We coded a factor FreeChoice such that an increment of 1 would always represent the difference between a pure FC reading and a pure uncertainty reading, as shown in Table 2. The two other factors, Scope and Speaker Knowledge received a treatment coding with **KS** and **NS** as baselines.

This analysis showed a significant interaction between Scope and FreeChoice ($\beta = -.18$, $\chi^2(1) = 10, p = .002$) and between Speaker Knowledge and FreeChoice ($\beta = .30, \chi^2(1) = 4.8, p = .028$). The triple interaction was not significant ($\beta = -0.07, \chi^2(1) = .7, p = .42$). A very puzzling fact however, is that we observed *more* FC readings in the **nKS** conditions than in the **KS** ones.

Since no theory predicts an effect in this direction, it suggests that the items we used in the **KS** context may have been less powerful at detecting FC choice readings. The **KS** target conditions are arguably difficult, as they require several steps of reasoning, and it may be that intermediate values reflect a difficulty in judging the sentence, rather than the realization that they are u.i. (our diagnosis for an uncertainty reading). This is partly confirmed by the fact that **UI** disjunction controls received much lower ratings than the **KS2** condition, although the former is only u.i. while the second is possibly ambiguous between an u.i. reading and a false one. We further tested the correlation between the average responses of each participant to the two types of KS targets. While we expected a negative correlation (both are u.i. without the FC reading, but **KS1** becomes true and **KS2** false if FC is derived), we actually observed a positive correlation ($\beta = .26, t = 2.1, p = .04$).

Given these results, we decided to run, as a post-hoc analysis, a model in which we replaced the problematic **KS2** condition with the **UI** disjunctive control. This would make the comparison between KS and NON-KS more minimal, and since the **UI** controls turned out to be very close to minimal acceptability, a FC reading would still correspond to a maximal difference. This analysis showed no interaction between FreeChoice and Speaker Knowledge ($\beta = -.17, \chi^2(1) = 2.3, p = .13$), and only a trend for an interaction between FreeChoice and Scope ($\beta = -.04, \chi^2(1) = 3.6, p = .06$), but revealed a significant negative triple interaction between FreeChoice, Scope and Speaker Knowledge ($\beta = -.20, \chi^2(1) = 6.6, p = .01$). This

	nKS		**KS1**	**KS2**	**UI**
	target	disj. ctrl	target	target	disj. ctrl
Truth value under FC reading: Truth value without FC reading:	false true	(true)	true u.i.	false u.i.	(u.i.)
Value of the FC factor used in statistical analyses	-0.5	0.5	0.5	-0.5	$-0.5^{\star}$

Table 2: Predicted truth value for each condition in Experiment 1 and definition of the factor used to detect FC readings in analyses. The idea was to compare two conditions which would be maximally distinct under a FC reading but equally acceptable otherwise. **UI** disjunction controls were only used in post-hoc analyses.

means that scope had a relatively small effect in **KS** contexts, but a significant one in **nKS** context. More precisely, the computed rates of FC were similar in **WS KS**, **NS KS** and **NS nKS**, but were reduced for **WS nKS** items. In Experiment 2, we improve on the design by simplifying the task, and replicate the result of this post-hoc analysis.

3 Experiment 2

3.1 Method

The design of Experiment 2 was substantially similar to that of Experiment 1, but we managed to simplify some key aspects. First, we made the participant the potential speaker, instead of the cook. This way, they only had to reason about their own knowledge, which was again manipulated by the context we gave them.Given the results of the previous experiment on u.i conditions, we decided to go on with a direct comparison and remove the second **KS** target. In the remaining **KS** condition, the context given to participants simply told them which grade the child was in. In the **nKS** condition, participants were provided partial information about which grade the child was in:

(9) a. **KS**: Mary is a 3rd grader.
b. **nKS**: We only know that Mary is not a 3rd grader

This experiment included the previous u.i.ând true disjunction controls (which were used as baselines for the **KS** and **nKS** targets, respectively), as well as a new false disjunction control, where the disjunctive statement was literally false. The number of repetitions for targets and controls was reduced to seven per condition to allow for more fillers (28 in total). The total number of items was thus 98, plus four training items. Fifty participants were recruited on MTurk; six were removed because of high error rates.

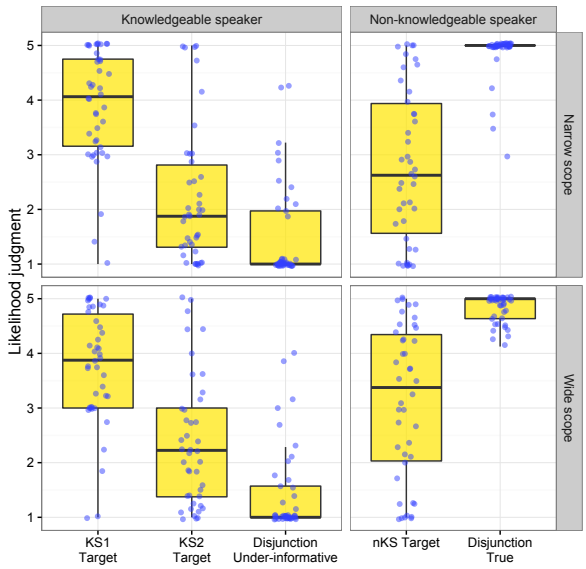
3.2 Results

The results for the relevant conditions are presented in Figure 1b. We ran an analysis similar to the post-hoc analysis of Experiment 1. We used the same sum-coded factor FreeChoice described in Table 2 (the only **KS** target left being equivalent to **KS1**). As previously, the two other factors, Scope and Speaker Knowledge received a treatment coding with **KS** and **NS** as baselines.

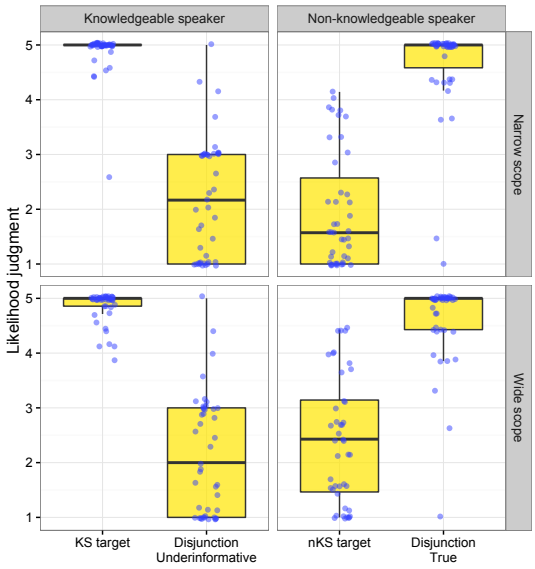
We observed a trend for an interaction between Scope and FC in the KS context ($\beta = -.01, \chi^2(1) = 3.1, p = .08$) but no difference between **nKS** and **KS** for **NS** sentences ($\beta = .01, \chi^2(1) = .6, p = .4$). We observed a highly significant triple interaction ($\beta = -0.25, \chi^2(1) = 7.6, p = .006$), thus replicating the post-hoc analysis of Experiment 1.

4 General discussion

We observed that the rate of FC was significantly reduced in the **WS nKS** condition compared to all others, although it appeared to be present in all conditions. This finding is in line with the predictions made by Aloni, 2016, since her account predicts that NS disjunction always give rise to a FC reading, while **WS** ones receive such a reading only if the speaker is knowledgeable. Since our **KS** conditions correspond to indisputable cases, our results match Aloni's predictions. Conversely, our findings contradict the predictions of the pragmatic approach in two respects. First, we observed FC readings in contexts where the speaker is not opinionated, which is unexpected if this is reading is an implicature. Second, all participants favored a FC interpretation for WS disjunctions (provided the speaker is knowledgeable). This contradicts judgments reported in Fox, 2007 which were generally accepted in the literature. In sum, our study suggests that Aloni's semantic account of FC ef-



(a) Experiment 1: FC leads to high scores in the KS1 condition and low scores in the KS2 condition. In principle, a non-FC reading should correspond to equal ratings for KS1, KS2 and the UI control condition. In nKS condition, FC corresponds to a low score in the target items



(b) Experiment 2: In the KS condition FC is manifested by high scores in the target items while in nKS condition it can be seen as a low score in the target items, with control conditions providing a non-FC baseline in each case.

Figure 1: Distribution of responses for each participant in Experiments 1 and 2 (averaged across items).

fect is in a better position than Fox's pragmatic one when it comes to interactions between scope and speaker knowledge.

These results have implications beyond the two analyses we discussed, and raise problems that any pragmatic theory must account for. While previous pragmatic theories have been able to deal with simple wide scope disjunction by reducing it to narrow scope via across-the-board movement (Meyer & Sauerland, 2016), this option is simply not available when the scope is marked overtly with 'either'. Hence, the results suggest that pragmatic approaches need to be updated in order to account for genuine wide scope FC, just like recent developments in semantic approaches allowed them to deal with NS disjunction.

References

Aloni, M. (2016). *FC disjunction in state-based semantics*. Handout.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax-pragmatics interface. In A. Belleti (Ed.), *Structures and beyond* (Vol. 3, pp. 39–103). Oxford: Oxford University Press.

Fox, D. (2007). Free choice disjunction and the Theory of scalar implicature. In U. Sauerland & P. Stateva (Eds.), *Presupposition and Implicature in Compositional Semantics* (pp. 71–120). New York, NY: Palgrave Macmillan.

Hawke, P. & Steinert-Threlkeld, S. (2016). Informational dynamics of epistemic possibility modals. *Synthese*, 1–34.

Kamp, H. (1973). Free choice permission. *Aristotelian Society Proceedings*, *74*, 57–74.

Kratzer, A. & Shimoyama, J. (2002). Indeterminate pronouns: the view from Japanese. In Y. Otsu (Ed.), *Third Tokyo conference on psycholinguistics* (pp. 1–25). Tokyo: Hituzi Syobo.

Larson, R. K. (1985). On the syntax of disjunction scope. *Natural Language and Linguistic Theory*, *3*(2), 217–264.

Meyer, M.-C. & Sauerland, U. (2016). *Covert across-the-board movement revisited*. Presentation at NELS 47 – UMass Amherst.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy*, *27*, 367–391.

Von Wright, G. H. (1968). An essay in deontic logic and the general theory of action. *Acta Philosophica Fennica*, *21*.

Zimmerman, T. E. (2000). Free choice disjunction and epistemic possibility. *Natural Language Semantics*, *8*, 255–290.