Reverse Sobel Sequences: What Is Being Cancelled Here?

David Krassnig
University of Konstanz

1 Introduction

Sobel sequences (SS) are sequences of conditionals of the pattern 'If ϕ, χ; but if $\phi \wedge \psi$, $\neg\chi$' (Stalnaker, 1968; Lewis, 1973), exemplified by (1a), and reverse Sobel sequences (rSS) follow the pattern pattern 'If $\phi \wedge \psi$, $\neg\chi$; but if ϕ, χ', exemplified by (1b). Especially the latter's general infelicity has played a pivotal role in the contemporary discussion on how conditionals should be modelled (Heim, 1994; von Fintel, 2001; Gillies, 2007; Klecha, 2014; Ippolito, 2020).

(1) a. If the USA threw its weapons into the sea tomorrow, there would be war; but if the USA and the other nuclear powers all threw their weapons into the sea tomorrow, there would be peace. (Lewis, 1973, p. 10)

b. If the USA and the other nuclear powers all threw their weapons into the sea tomorrow, there would be peace; #but if the USA threw its weapons into the sea tomorrow, there would be war. (Heim, 1994)

Particularly complicating has been the discovery of felicitous rSS (Moss, 2012):

(2) Suppose John and Mary are our mutual friends. John was going to ask Mary to marry him, but didn't at the last minute. I know Mary much better than you do, and you ask me whether Mary might have said yes if John had proposed. I tell you that I swore to Mary that I would never tell anyone that information, which means that strictly speaking, I cannot answer your question. But I say that I will go so far as to tell you two facts:

a. If John had proposed to Mary and she had said yes, he would have been really happy.

b. But if John had proposed, he would have been really unhappy.

Whilst Moss (2012) originally proposed that rSSs are rendered infelicitous by the epistemic irresponsibility of their utterance, subsequent research showed that this is far from the entire story (Klecha, 2014; Lewis, 2018; Krassnig, 2020).

In this paper we examine one recent proposal for the infelicity of rSSs and propose an alternative account based upon a variably-strict semantics, briefly explained in the next subsection, that encompasses more of the known empirical data listed below.

The following factors have been identified in the literature as determining the (in)felicity of rSS: (i) For felicity, the utterance of the ϕ-conditional must be considered epistemically responsible despite the utterance of the $\phi \wedge \psi$-conditional (Moss, 2012). (ii) If there is a causal relation between ϕ and ψ, any rSS is infelicitous. (iii) There must be some type of contrastive stress in the antecedent of the second condtional to allow felicity (Klecha, 2014; Krassnig, 2020); and (iv) if there is no overtly different lexical item in the antecedent, the auxiliary verb must be contrastively stressed to allow for the possibility of felicity (Krassnig, 2020).

1.1 Variably-Strict Semantics

Variably-strict semantics restrict the domain quantified over by conditionals to the maximally similar antecedent world and check whether all of these worlds are also consequent worlds:

(3) For all contexts c, 'If ϕ, ψ' is true at w in c iff all the closest ϕ-worlds to w are ψ-worlds, where closeness is determined by similarity.

This system is visually demonstrated by Figure 1. Similarity is determined by how many cause-initial deviances the respective world introduces into to the evaluation world by comparison.

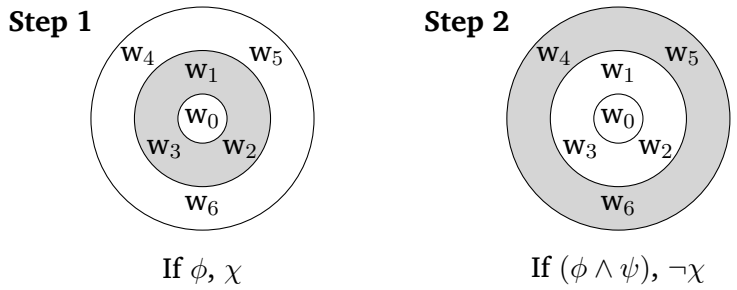


Figure 1: Domains of quantification for SSs according to Stalnaker (1968) and Lewis's (1973) variably-strict conditional analyses. For all w_n-worlds: If $n \geqslant 1$, then $\phi = 1$ holds true for w_n, and if $n \geqslant 4$, then $\psi = 1$ holds true for w_n also. ϕ does not precede ψ on a causal event chain.

2 Ippolito's (2020) Specificity-Condition-Based Account

Ippolito (2020) attempts to establish a connection between hereto two independent phenomena with similar effects: the generally unidirectional felicity of disjunctive sentences such as the ones in (2) and the generally unidirectional felicity of SS (original idea due Singh, 2008).

(4) a. John will eat some cookies or he will eat all the cookies.
b. # John will eat all the cookies or he will eat some cookies.

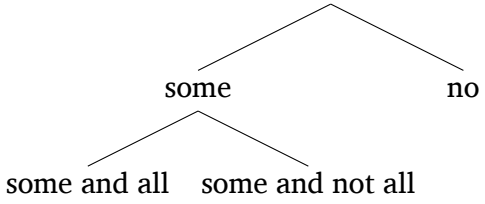
(Ippolito, 2020, p. 634)

To explain how Ippolito (2020) derives the infelicity of (1b) and (4b), we must first explain how Ippolito (2020) handles the generation of alternatives, as this is crucial to her account, as will soon become apparent. Ippolito (2020) takes alternatives (in the sense of Rooth (1992)) to be possible answers to some question under discussion (QUD; in the sense of Büring (2003); Roberts (1996)), where the alternatives generated by some sentence S with the form $[_S \ldots \alpha_F \ldots]$ make up a structured alternative set T_{A_α} s.t. the following conditions are fulfilled:

(5) T_A is a well-formed structured set of alternatives iff the following conditions are met:
a. *Strength:* For any two alternatives $\alpha, \beta \in A$, β is the daughter of α in T_A if $[\![\beta]\!] \subseteq [\![\alpha]\!]$
b. *Disjointness:* For any $\beta_1, \beta_2 \in A$, if β_1, β_2 are sisters in T_A, then $[\![\beta_1]\!] \cap [\![\beta_2]\!] = \emptyset$
c. *Exhaustivity:* For any alternative α with daughters $\beta_1, \ldots, \beta_n$ in T_A, $[\![\beta_1]\!] \cup [\![\beta_2]\!] = [\![\alpha]\!]$

As such, a focused *some*$_F$ would yield the following structure of alternatives:[1]

(6)



[1]Note that Ippolito (2020) crucially assumes that the focus alternatives of *some* are not limited to *some* and *all*, but include *no* and *some and not all* as well.

Ippolito (2020) argues that a sequence of two focus-containing sentences that answer the same QUD are only felicitous if they are maximally informative with respect to each other, i.e., if they are at the same level of specificity within the structured set of alternatives:

(7) *Specificity Condition:* A sequence $\Sigma < [_{S_i} \dots \alpha_F \dots], [_{S_j} \dots \beta_F \dots] >$, s.t. both S_i and S_j are answers to the same QUD and $\beta \in T_{A_\alpha}$, is felicitous if either:
   a. α or β is the only node on its branch in T_{A_α}, or
   b. α and β are dominated by he same number of nodes in T_{A_α}

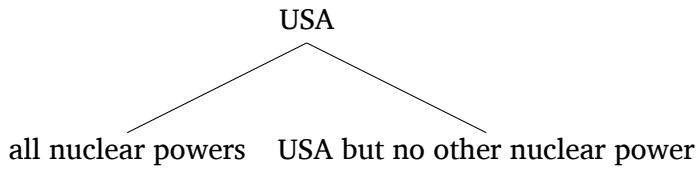
If a sequence violates the specificity condition, the weaker item in the sequence is covertly strengthened iff (i) the covertly strengthened reading would then satisfy the specificity condition and (ii) the covertly strengthened reading is neither equal to the sibling or mother node of the preceding utterance in the structured set of alternatives (for reasons of economy; see Ippolito, 2020, p. 643). If the second condition was not fulfilled, covert strengthening would be illicit and the weaker item would have to be overtly strengthened to derive felicity. In the case of (4a), Ippolito (2020) argues that the first disjunct is strengthened via covert exhaustification (Chierchia et al., 2012), because there is no preceding item in the sequence and exhaustification would satisfy the specificity condition:

(8) a. exh(John will eat some$_{\mathrm{F}}$ cookies) or he will eat all the cookies
   b. $exh(C)(\phi)(w) = 1$ iff $\phi(w) = 1$ and $\forall \psi \in C : \psi(w) = 1 \rightarrow (\phi \rightarrow \psi)$

In the case of (4b), felicity cannot be derived because *John will eat some and not all cookies* is a sibling of the preceding item in the sequence in the structured set of alternatives.

Ippolito (2020) argues that the same mechanism derives the infelicity of rSSs, answering the conditional QUD (CQUD) 'if what, χ?'. The only difference is the mechanism responsible for the covert strengthening: Whilst exhaustification was used for disjunctive sequences, Ippolito (2020) argues that the way similarity restricts the worlds quantified over in conditionals is akin to a covert strengthening mechanism that is bound by the same conditions. After all, if the closest ϕ-worlds are more similar to w_0 than the closest $\phi \wedge \psi$-worlds, it is similar to strengthening ϕ to $\phi \wedge \neg\psi$, since there are no ψ-worlds in the ϕ-conditional's domain of quantification. Furthermore, due to the nature of rSSs, the covert similarity-'strengthening' of the ϕ-conditional would typically correspond to a sister node of the preceding $\phi \wedge \psi$-conditional. This is demonstrated below, using the focused elements from (1b) as an example:

(9)



rSSs would therefore typically violate the economy constraints placed upon covert strengthening by Ippolito (2020), preventing its usage, thereby violating the specificity condition. As such, to derive a felicitous rSS, the structured set of alternatives generated must be reordered s.t. the sequence no longer violates the specificity condition. One way to accomplish this is to exclude some of the possible answers to the CQUD. In (2), for example, we can exclude all answers where Mary accepted John's proposal for the ϕ-conditional due to context, after having entertained the hypothetical possibility of him having done so for the $\phi \wedge \psi$-conditional. This way, we would end up with a structured set of alternatives consisting of only two possibilities: *John proposed and Mary said no* and *John didn't propose*. This way, the specificity condition may be fulfilled, resulting in a felicitous rSS (for details, see Ippolito, 2020, p. 663).

With this Ippolito (2020) may account for two of our known empirical pieces of data: That rSSs are felicitous if the $\phi \wedge \psi$-conditional's utterance may be considered epistemically responsible (which is only possible if we have a structured set of alternatives where some alternative answers have been excluded) and the presence of stress, in general, in the antecedent (since Ippolito (2020) requires focus accent to identify the CQUD). The remaining data, however, is not explicitly accounted for as of yet.

We have two possible points of detraction from this model: First, similarity does not always act akin to strengthening. One of the unaccounted for pieces of data, for example, are rSSs where ϕ and ψ are causally related. Adopting Bennett's (2003) view on world similarity, ψ would not further decrease the similarity of ϕ-worlds if ϕ precedes ψ on some causal chain of events. As such, for causal rSSs, there is no shift in similarity from the $\phi \wedge \psi$-conditional to the ϕ-conditional. It should, therefore, not be subject to restraints on covert strengthening and therefore not subjecto to Ippolito's (2020) reasoning. Their universal infelicity may still easily be derived, however, by claiming that their contradictory claims over partially the same worlds is the source of their infelicity (Klecha, 2014). As such, this is more of an amendment to Ippolito (2020) than an actual detraction. Our second point would be the difference in nature between covert exhaustification and similarity-based world restriction. The former is merely a conversational implicature and therefore cancellable. As such, it is not necessarily surprising that its generation may be prevented, even if this results in infelicity. The similarity-based restriction of worlds, on the other hand, is considered to be an integral part of the literal and semantic meaning of conditionals. It is therefore not as apparent to us how a semantic meaning may be cancelled or prevented via pragmatics—only to lead to unnecessary infelicity.

3 Our Contrastive-Stress-Based Account

With this, we propose our own account: We argue that all properties of rSSs revolve around the meaning of contrastive stress in the antecedent. More specifically, the meaning of contrastively stressed auxiliary verbs in rSSs. We do this by adopting the following pre-existing assumptions: (i) The antecedent of a conditional sets the current aboutness topic (Ebert et al., 2008); (ii) *would* is sensitive to modal subordination (Klecha, 2011, 2014); (iii) the focus value for pro-forms may be a set of identity functions over alternative domains (Jacobson, 2004); and (iv) differences to the actual world that causally stem from another initial change do not further decrease the similarity of the world in question (Bennett, 2003; Arregui, 2009).

We would argue the following: The first conditional's antecedent sets the current aboutness topic (Ebert et al., 2008). If this is followed by another conditional with a topically compatible antecedent, a subordinating discourse relation is established between the two such that the latter conditional is considered modally subordinate to the former. We posit that the topics generated by ϕ-antecedents and $\phi \wedge \psi$-antecedents are always considered compatible and thereby subordinating towards one another. This way, a rSS 'If $\phi \wedge \psi$, $\neg\chi$; but if ϕ, χ' is analysed as 'If $\phi \wedge \psi$, $\neg\chi$; but if $\phi \wedge (\phi \wedge \psi)$, χ', resulting in directly contradictory claims over the same sets of worlds. We would argue that this is the source for the general infelicity of rSSs (as already partially argued for by Klecha (2014)).

The only way to escape modal subordination and, thereby, infelicity is to indicate that the aboutness topic changes between conditionals. To do this, for rSSs, contrastive topic is required. Contrastive topic requires two lexical items of different semantic values: Either we have some overtly different lexical items or overtly identical lexical items with covertly different semantic values. The former case would be rather straightforward: If an overtly different lexical item is contrastively stressed, this invokes the exhaustification operator, thereby negating all non-entailed alternatives. For rSSs, this would correspond to the following reading: 'If $\phi \wedge \psi$, $\neg\chi$;

but if $\phi \wedge \neg\psi$, χ.' As this topic is ostensibly incompatible with $\phi \wedge \psi$, the modally subordinate reading is prevented. For overtly identical lexical items with differing semantic values, our options are limited and more complicated: We would posit that the auxiliary verb, which appears to be obligatorily targeted in such cases, is our only candidate. But to determine this, we need to determine which part of meaning of the auxiliary verb is targeted.

We argue that the auxiliary verb is obligatorily targeted due to our attempt to target the tense, aspect, and mood (TAM) morphology encoded by it. The TAM morphology, in turn, we target due to its connection to how the worlds quantified over by the conditional are selected and which properties they must have (Palmer, 1986; Iatridou, 2000; Arregui, 2009; Romero, 2014; Schulz, 2014, amongst others). In fact, we posit that the contrastively stressed TAM morphology behaves exactly like a pro-form bound by the domain of the closest antecedent worlds. This yields the question of how contrastively stressed pro-forms function. Contrastively stressed pro-forms are only felicitous if they are bound by disjoint domains:

(10) a. Every fourth grade boy$_i$ called his$_i$ mother, but no *fifth* grade boy$_j$ called *his*$_j$ mother.

b. * I expected every student$_i$ to call his$_i$ father, but only every *young* student$_j$ called *his*$_j$ father. (Sauerland, 1998, p. 206)

We follow Jacobson (2000, 2004) in assuming that pro-forms are identity functions and that bound pro-forms are partial identity functions restricted to members of its domain:

(11) a. every third-grader $[\lambda x.\textsc{love}(x, \textsc{the-mother-of}(\textsc{identity}_{\textsc{3rd-graders}}(x)))]$

b. every fourth-grader $[\lambda x.\textsc{love}(x, \textsc{the-mother-of}(\textsc{identity}_{\textsc{4th-graders}}(x)))]$

Jacobson (2000, 2004) then derives the contrastive stress by contrastively stressing the two differing partial identity functions (one that only ranges over the third graders and one that only ranges over the fourth graders and would return an undefined value otherwise). The contrastive stress is only felicitous if the two identity functions are disjoint in their domains.

Since we consider the auxiliary verb to be identical in meaning to a pro-form bound by the domain of the closest antecedent worlds, we consider its meaning to be equal to a partial identity function of type $<s, s>$ restricted to the aforementioned domain:

(12) a. If $[\lambda w.\phi(\textsc{identity}_{\textsc{closest-}\phi \wedge \psi}(w)) \wedge \psi(\textsc{identity}_{\textsc{closest-}\phi \wedge \psi}(w))]$, then …

b. If $[\lambda w.\phi(\textsc{identity}_{\textsc{closest-}\phi}(w))]$, then …

The contrast is then derived by comparing the two different partial identity functions, whose contrast is only felicitous if their ranges are disjoint to one another. If the contrast is felicitous, the contrastive topic is successful in preventing the modally subordinate reading, yielding felicitous rSSs. For counterfactuals, in a variably-strict semantics, the domains of the closest ϕ-worlds and of the closest $\phi \wedge \psi$-worlds are disjoint iff ϕ does not precede ψ on any causal chain of events (Bennett, 2003; Arregui, 2009).

With these assumptions, we can derive all of the known empirical data: (i) Contrastive stress is required to prevent modal subordination; (ii) contrastively stressed auxiliary verbs enforce disjoint domains of quantification, preventing modal subordination; (iii) rSSs are felicitous if uttering the ϕ-conditional is epistemically responsible despite the $\phi \wedge \psi$-conditional's previous utterance, since this may only ever be the case if we can be assured that the possibility of ψ does not arise from the assumption of ϕ, which is only the case if both domains of quantification are disjoint; and (iv) causal rSSs are always infelicitous, since, if ϕ precedes ψ on some causal chain of events, the closest $\phi \wedge \psi$-worlds would be a mere subset of the closest ϕ-worlds (Bennett, 2003; Arregui, 2009; Klecha, 2014), ensuring that the disjoint domain condition may never be fulfilled. As such, we would argue that the cancellation of modal subordination enables rSS felicity, rather than the cancellation of similarity-'strengthening' causing the infelicity of rSS.

References

Arregui, Ana. 2009. On similarity in counterfactuals. *L&P* 32(3). 245–278.

Bennett, Jonathan. 2003. *A philosophical guide to conditionals*. Oxford: Oxford University Press.

Büring, Daniel. 2003. On D-trees, beans, and B-accents. *L&P* 26. 511–545.

Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2012. The grammatical view of scalar implicatures and the rela-tionship between semantics and pragmatics. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *An international handbook of natural language meaning*, vol. 3, 2297–2332. Berlin: Mouton de Gruyter.

Ebert, Christian, Cornelia Endriss & Stefan Hinterwimmer. 2008. Topics as speech acts: An analysis of conditionals. In Natasha Abner & Jason Bishop (eds.), *Proceedings of the 27th WCCFL*, 132–140. Somerville, MA: Cascadilla Proceedings Project.

von Fintel, Kai. 2001. Counterfactuals in a dynamic context. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 123–152. Cambridge: MIT Press.

Gillies, Anthony S. 2007. Counterfactual scorekeeping. *L&P* 30(3). 329–360.

Heim, Irene. 1994. Unpublished MIT seminar on conditionals.

Iatridou, Sabine. 2000. The grammatical ingredients of counterfactuality. *Linguistic Inquiry* 31(2). 231–270.

Ippolito, Michela. 2020. Varieties of sobel sequences. *L&P* 43. 633–671.

Jacobson, Pauline. 2000. Paychecks, stress, and variable-free semantics. In Jackson Brendan & Tanya Matthews (eds.), *Proceedings of SALT X*, 65–82. Ithaca, New York: Cornell University.

Jacobson, Pauline. 2004. Kennedy's puzzle: What I'm named or who I am. In R. Young (ed.), *Proceedings of SALT XIV*, 145–162.

Klecha, Peter. 2011. Optional and obligatory modal subordination. In Ingo Reich, Eva Horch & Dennis Pauly (eds.), *Proceedings of SuB 15*, 365–379.

Klecha, Peter. 2014. *Bridging the divide: Scalarity and modality*: Uni of Chicago dissertation.

Krassnig, David. 2020. Reverse Sobel sequences and the dissimilarity of antecedent worlds. In N. Franke, N. Kompa, M. Liu, J. Müller & J. Schwab (eds.), *Proceedings of SuB 24*, 447–463.

Lewis, David Kellog. 1973. *Counterfactuals*. Malden: Blackwell.

Lewis, Karen. 2018. Counterfactual discourse in context. *Noûs* 52(3). 481–507.

Moss, Sarah. 2012. On the pragmatics of counterfactuals. *Noûs* 46(3). 561–586.

Palmer, Frank Robert. 1986. *Mood and modality*. Cambridge: Cambridge University Press.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. In J. Yoon & A. Kathol (eds.), *Papers in semantics*, vol. 49 OSU Working Papers in Linguistics, 91–136. Ohio State University, Department of Linguistics.

Romero, Maribel. 2014. 'Fake tense' in counterfactuals: A temporal remoteness approach. In Luka Crnič & Uli Sauerland (eds.), *The art and craft of semantics: A festschrift for irene heim*, vol. 2, 47–63. MITWPL 71.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116.

Sauerland, Uli. 1998. *The meaning of chains*. Cambridge: MIT dissertation.

Schulz, Katrin. 2014. Fake tense in conditional sentences: A modal approach. *Natural Language Semantics* 22. 117–144.

Singh, Raj. 2008. On the interpretation of disjunction: Asymmetric, incremental, and eager for inconsistency. *Linguistics and Philosophy* 31(2). 245–260.

Stalnaker, Robert Culp. 1968. A theory of conditionals. In N. Rescher (ed.), *Studies in logical theory*, 98–112. Oxford: Basil Blackwell Publishers.