
Weakness of will

Julie Goncharov

University of Göttingen

1 Introduction

In some scenarios, sentences like in (1) have a clear *weakness of will* inference which states that the speaker was initially resolved or committed to one course of action, but acted contrary to her resolution. Such scenarios would describe a person on a weight loss program not being able to resist delicious dessert in (1)a and a lover overwhelmed with feelings towards the past relationship in (1)b.

- (1) a. I didn't want to eat this cake, but I did. (weakness of will)
b. I didn't want to call my ex, but I did. (weakness of will)

Similar configurations in (2), however, do not give rise to the same weakness of will inference, even when (2)a describes a loving husband trying to communicate to his wife that her cooking skills need some improvement without offending her and (2)b describes an expert carefully examining an antique vase so as to not break it.

- (2) a. I didn't want to offend you, but I did. (*weakness of will)
b. I didn't want to break the vase, but I did. (*weakness of will)

In this paper, we address the following question: what is the difference between predicates like *eat* and *call*, on the one hand, and *offend* and *break*, on the other hand, that is responsible for the presence/absence of weakness of will inference? The answer we arrive at is that the former, but not the latter, express actions that can be fully controlled by the agent. We argue that weakness of will can be used as a linguistic test to distinguish intentional versus accidental actions. We see that the relevant distinction between intentional and accidental actions cannot be detected by other linguistic tests, such as modification by *(un)intentionally* or rationale clauses introduced by *in order to* (e.g., Jackendoff 1972, 1995; Farkas 1988, 1992). The intentional/accidental distinction is crucial for analysing many phenomena that have recently received a significant amount of attention in the literature, such as generalized subject obviation (Oikonomou 2016; Stegovec 2017, 2019; Kaufmann 2019; Szabolcsi 2010, to appear, a.o.), licensing of polarity sensitive items (Szabolcsi 2010; Zu 2018; Goncharov 2020a, a.o.), and aspect choice in Slavic (Despić 2020; Goncharov 2020b, a.o.).

2 Preliminaries

Weakness of will is a topic widely discussed in the philosophy of action. Starting from Aristotle, weakness of will has been intimately connected to *akrasia*, i.e., acting against one's best judgments (Aristotle, *Ethica Nicomachea* in Aristotle 1941). In this paper, we take weakness of will in its narrow sense as it is understood in Holton 1999, 2009 or what Mele (1987, 1995, 2010) calls 'unorthodox akrastic actions'. To see the difference between weakness of will in the narrow sense and akrastic actions in its traditional sense, consider an example from Holton (2009):¹

“Christabel, an unmarried Victorian lady, has decided to embark on an affair that she knows will be disastrous. It will ruin her reputation, and quite probably leave her

¹This scenario is adopted from A. S. Byatt's novel *Possession*, 1989.

pregnant. Moreover, she considers it morally wrong. So she thinks it not the best option on either moral or prudential grounds. Nevertheless, she has resolved to go ahead with it. However, at the very last moment she pulls out: not because of a rational reconsideration of the pros and cons, but because she simply loses her nerve.” (Holton 2009, p. 84)

In this scenario, Christabel’s going ahead with the affair would be acting against her best judgments, i.e., akrasia in its traditional sense. But Christabel’s not going ahead with the affair is weakness of will in the narrow sense that interests us here. Weakness of will in this sense is associated with acting against one’s previous decision or executive commitment. Following Holton (1999, 2009), we take weakness of will to be *unjustified reconsideration of one’s decision or executive commitment*.²

3 Implementation

The scenarios for examples in (1) and (2) are constructed in a way to ensure neg-raising interpretation of *want*. That is to say, (1)a is to be interpreted as *I wanted not to eat the cake* and (2)a as *I wanted not to offend my wife*. This neg-raising interpretation aligns with the so-called ‘effective-preference’ reading of *want* as the purely psychological reading of *want* is incompatible with neg-raising (e.g., *I don’t want to kiss Mary* $\not\leftrightarrow$ *I want not to kiss Mary*, see Davis 1984; Condoravdi & Lauer 2012).

To account for weakness of will, we begin with the observation that effective-preference reading of *want* is what is called ‘decision’ or ‘executive commitment’ in the philosophical literature on the topic. Let us say that in sentences with *want*, ϕ represents the proposition that describes the action (i.e., the prejacent) and ψ stands for the complex proposition ‘the agent acts so as to bring it about that ϕ ’. Then, we say that in *want*-sentences, we have:

- (3) Decision/executive commitment:
- a. $\Box^{B_x}\psi$ when ‘ x want ϕ ’ is uttered
 - b. $\Box^{B_x}\neg\psi$ when ‘ x not want ϕ ’ is uttered (under neg-raising interpretation)

Additionally, we propose that actions can differ with respect to whether they are *fully controlled* by the agent or part of their execution is left to chance. For the moment, we focus on the actions that are initiated intentionally, i.e., the conditions in (3) obtain. We say that an action is interpreted as *controlled* when the agent x of the action believes that if she acts so as to bring it about that ϕ , the state of affairs described by ϕ obtains and similarly for $\neg\phi$. An action is interpreted as *non-controlled* when the negation of the controlled condition holds. The controlled and non-controlled conditions are formalized as follows:

- (4) a. Controlled actions: $\Box^{B_x}((\psi \rightarrow \phi) \wedge (\neg\psi \rightarrow \neg\phi))$
 b. Non-controlled actions: $\neg\Box^{B_x}((\psi \rightarrow \phi) \wedge (\neg\psi \rightarrow \neg\phi))$

With these conditions in mind, it is easy to see why weakness of will surfaces with controlled actions, but not with non-controlled actions. Let us first take up controlled actions like *eat* in (1)a. As shown in (5), for the belief state of the speaker (who is the attitude holder and the agent of the action) to be consistent, there should be an unjustified revision of the agents decision not to eat the cake. This revision (or unjustified reconsideration) surfaces as the weakness of will inference.

- (5) a. I didn’t want to eat this cake, but I did.
 b. $\Box^{B_x}\neg\psi \wedge \Box^{B_x}(\neg\psi \rightarrow \neg\phi)$, but $\Box^{B_x}\phi$
 Unjustified reconsideration: $\Box^{B_x}\psi \wedge \Box^{B_x}(\psi \rightarrow \phi)$

²For reasons of space, we leave out the discussion of what counts as ‘unjustified reconsideration’, see Holton 2009.

With non-controlled action like *offend*, the reconsideration of one’s executive commitment is not necessary as the belief state of the attitude holder (who is also the agent of the action) is compatible with the continuation ...*but I did*:

- (6) a. I didn’t want to offend you, but I did.
 b. $\Box^{B_x} \neg \psi \wedge \Diamond^{B_x} (\neg \psi \wedge \phi)$, but $\Box^{B_x} \phi$

It is trivial to see that in object-control configurations like *I didn’t want John to eat this case/offend you, but he did*, no weakness of will is predicted.

4 Application

In the recent literature, the interpretation of actions as intentional versus accidental has been argued to be a determining factor in explaining phenomena like exceptions to generalized subject obviation (e.g., Kaufmann 2019; Szabolcsi 2010, to appear and references therein), licensing of polarity sensitive items (e.g., Szabolcsi 2010; Zu 2018; Goncharov 2020a), aspect choice in Slavic (e.g., Despić 2020; Goncharov 2020b), and others. However, the relevant *linguistic* distinction between intentional and accidental actions is not well captured in this literature. In particular, two linguistic tests have been proposed to tell apart intentional and accidental interpretations of actions: (i) modification by *(un)intentionally* and (ii) modification by rationale clauses introduced by *in order to* (e.g., Jackendoff 1972, 1995; Farkas 1988, 1992). We see below that neither of these tests provides an adequate way for distinguishing between intentional versus accidental interpretation that is relevant for the phenomena above, whereas the weakness of will test makes the correct distinction.³

To demonstrate this point, we take one particular phenomena sensitive to the interpretation of an action as intentional versus accidental, namely anti-licensing of Positive Polarity Items (PPIs) in the complement of *not want*.

Szabolcsi (2004) observes that anti-licensing of PPIs formed with *some* in the infinitival complement of *not want* is sensitive to the interpretation of an action in the complement clause. When the action is intentional (as in the case of *call* and *eat*), *some* cannot have a narrow scope interpretation, see (7)a. But with accidental actions (like *offend* or *break*), *some* can be interpreted under negation, see (7)b. Sensitivity of *some* PPIs to the interpretation of an action as (un)intentional is also attested in Hebrew, Hungarian, Polish, Romanian, and Russian.

- (7) a. I don’t want to call someone/eat something. (some>not, *not>some)
 b. I don’t want to offend someone/break something. (some>not, not>some)

To predict the behaviour of *some* for other verbs, we’d like to have a linguistic test that differentiates *call/eat* from *offend/break*. Let us begin with the modification by *(un)intentionally* test. This test can be dismissed as unreliable on the following grounds: First, although *(un)intentionally* can tell apart actions for which the agent is responsible, as in (8)a,b from actions for which he is not, as in (8)c,d, these adverbials are also incompatible with actions of the same kind. As illustrated in (9), *intentionally* is infelicitous with clearly intentional actions (because the modification is superfluous).⁴

- (8) a. John hurt Mary intentionally.
 b. John fell off the ladder intentionally.
 c. #John resembles his father intentionally.
 d. #John is tall intentionally.

³This is not to deny that the tests are capturing other aspects of the meaning of the action in question.

⁴*(Un)intentionally* and its kin are focus sensitive. We put this point aside here.

- (9) a. #Mary baked the cake intentionally.
 b. #John wrote the novel intentionally.

Second, modification by (*un*)*intentionally* is very context-dependent and both *intentionally* and *unintentionally* can modify actions like *call*, *eat* and *offend*, *break*. The examples in (10)a,b show that *unintentionally* can modify verbs that normally express intentional actions, such as *call*, *eat*. The examples in (10)c,d illustrate that verbs describing stereotypically unintentional actions, such as *lose*, *win*, can be modified by *intentionally* when a correct context is provided.

- (10) a. If an organization unintentionally calls a mobile number without consent, it is safeguarded against prosecution.
 b. Would you still be a vegan if you unintentionally ate an animal product?
 c. I intentionally lost a few chips so I would have less than everyone else.
 d. Nate has intentionally won a few games, I'm still not brave enough to do that as I don't like to see my little buddy upset and am a born people pleaser.

Let us now look at modification by rationale clauses introduced by *in order to*. The examples in (11) show that rationale clauses are only compatible with actions that can be intentionally initiated by the agent:

- (11) a. John went home in order to see his mother.
 b. John bought a car in order to get to work quicker.
 c. #John grew taller in order to ...
 d. #John is descended from royalty in order to ... (Jackendoff 1995, 220)

Because the verbs we are interested in (*call*, *eat*, *offend*, *break*) all express actions that can be initiated intentionally, modification by *in order to* cannot distinguish between them:

- (12) a. After a few minutes, the driver of the vehicle called the police in order to file a report.
 b. Eve ate the fruit in order to fulfill desires that God had already offered to meet.
 c. There can also be times when he will step on other people's toes and offend some sensibilities in order to achieve his goal.
 d. Thomas broke the handle in order to enter the room.

Given these facts, we highlight that there is a distinction between actions that are *initiated intentionally* and *intentional actions*, which we understand as actions that are initiated intentionally and are controlled by the agent. Actions that are initiated intentionally can be tested using modification by *in order to* clauses and (*un*)*intentionally*. The weakness of will test adds testing for the controllability component which eventually tests for intentional actions. Below we see that it is the latter distinction that is needed to account for the PPI facts.

Let us assume the proposal in Zeijlstra to appear that *some* PPIs are licensed when the Non-entailment of non-existence condition in (13) is met, which is corroborated by the well-known fact that *some* is not anti-licensed in non-anti-additive downward-entailing environments, such as *At most five men called someone* (at most < some).

- (13) Non-entailment of non-existence condition: the use of a PPI is not felicitous iff the proposition in which it appears entails non-existence of a referent satisfying its description.

In sentences with intentional actions the use of *some* is infelicitous because the control condition renders the epistemic modal base of *want* anti-additive, see (14). On the other hand, the non-control condition allows (13) to be met, thus *some* is not anti-licensed, see (15).

- (14) a. I don't want to call someone. (*not>some)
 b. $\Box^{B_x} \neg \psi \wedge \Box^{B_x} (\neg \psi \rightarrow \neg \phi) \rightsquigarrow \Box^{B_x} \neg \phi$
- (15) a. I don't want to offend someone. (not<some)
 b. $\Box^{B_x} \neg \psi \wedge \Diamond^{B_x} (\neg \psi \wedge \phi)$

We will show in the paper that a similar analysis can be given to other phenomena sensitive to the interpretation of an action as intentional versus accidental.

5 Conclusion

The concept of intentions and intentional actions has received a lot of attention in philosophy starting from Aristotle. Since Aristotle, philosophical reflections on the nature of intentions and intentional actions have revolved around two key components: *desire* and *belief* (Anscombe 1957; Davidson 1963, 1980, 2001; Bratman 1987, 1999; Velleman 1989; Raz 2011, a.o.). In more recent literature, a third component - *control*, “man’s own power” - has been added. When an action is beyond one’s abilities or is performed under coercion, such an action is not intentional, although the outcome may be highly preferable and foreseen (Mele and Moser, 1994, a.o.).

Philosophical inquiries focus on understanding the folk concept of intentions and their insights cannot be straightforwardly carried to linguistics which should concern itself only with those aspects of intentions and intentional actions that are relevant for purposes of linguistic communication. The investigation in this paper is one of the first steps towards constructing philosophically informed linguistic notion of intentional actions.

References

- Anscombe, G. E. M. 1957. *Intention*. Cambridge, London: Harvard University Press.
- Aristotle. 1941. *The basic works of Aristotle*,. New York: Random House.
- Bratman, Michael. 1987. *Intentions, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press.
- Bratman, Michael. 1999.. 1999. *Faces of intention*. Cambridge: Cambridge University Press.
- Condoravdi, Cleo & Sven Lauer. 2012. Imperatives: meaning and illocutionary force and illocutionary force. In Christopher Pi non (ed.), *Empirical Issues in Syntax and Semantics*, vol. 9, 37–58.
- Davidson, Donald . 1963. Actions, reasons, and causes. *The journal of Philosophy* 60(23). 685–700.
- Davidson, Donald. 1980. *Essays on actions and events*. Oxford: Clarendon Press.
- Davidson, Donald. 2001. *Essays on Actions and Events*. Oxford: Oxford University Press.
- Davis, Wayne A. 1984. The two senses of desire. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 45(2). 181–195.
- Despić, Miloje. 2020. Negative imperatives and aspect: a phase-based approach. In P. Kosta & T. Radeva-Bork (eds.), *Current Developments in Slavic Linguistics: Twenty Years After*, Bern: Peter Lang.
- Farkas, Donka. 1988. On obligatory control. *Linguistics and Philosophy* 11. 27–58.
- Farkas, Donka F. 1992. On obviation. In Ivan A. Sag & Anna Szabolcsi (eds.), *Lexical Matters*, chap. 4, 85–110. Stanford, Calif.: Center for the Study of Language and Information.
- Goncharov, Julie. 2020a. The dynamic presupposition of ‘want’. *Proceedings of SALT 30* .
- Goncharov, Julie. 2020b. Whom to oblige? In Franc Marušič Marušič, Petra Mišmaš & Rok Žaucer (eds.), *Advances in formal Slavic linguistics 2017*, 51–73. Berlin: Language Science Press.

- Holton, Richard. 1999. Intention and weakness of will. *The Journal of Philosophy* 96(5). 241–262.
- Holton, Richard. 2009. *Willing, Wanting, Waiting*. Oxford: Clarendon Press.
- Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge: MIT Press.
- Jackendoff, Ray. 1995. The conceptual structure of intending and volitional action. In Héctor Campos & Paula Kempchinsky (eds.), *Evolution and Revolution in Linguistic Theory*, 198–227. Washington, D.C.: Georgetown University Press.
- Kaufmann, Magdalena. 2019. Who controls who (or what). *Proceedings of SALT 29* 636–664.
- Mele, Alfred R. 1987. *Irrationality*. New York: Oxford University Press.
- Mele, Alfred R. 1995. *Autonomous agents*. New York: Oxford University Press.
- Mele, Alfred R. 2010. Weakness of will and akrasia. *Philosophical Studies* 150(3). 391–404.
- Oikonomou, Despina. 2016. *Covert modals in root contexts*: MIT dissertation.
- Raz, Joseph. 2011. *From Normativity to Responsibility*. Oxford: Oxford University Press.
- Stegovec, Adrian. 2017. !? (where's the ban on imperative questions?). *SALT 27* 153–172.
- Stegovec, Adrian. 2019. Perspectival control and obviation in directive clauses. *Natural Language Semantics* 27(1). 47–94.
- Szabolcsi, Anna. 2004. Positive polarity – negative polarity. *Natural Language and Linguistic Theory* 22. 409–452.
- Szabolcsi, Anna. 2010. Infinitives vs. subjunctives: What do we learn from obviation and from exemptions from obviation? URL <http://goo.gl/c16gIP> ms. New York University.
- Szabolcsi, Anna. to appear. Obviation in hungarian: What is its scope, and is it due to competition? *Glossa* .
- Velleman, J. David. 1989. *Practical reflection*. Princeton: Princeton University Press.
- Zeijlstra, Hedde. to appear. *Negation and negative dependencies*. Oxford University Press.
- Zu, Vera. 2018. *Discourse participants and the structural representation of the context*: NYU dissertation.