

Sociophonetics, semantics, and intention

Eric K. Acton

Eastern Michigan University

Overview. As distinct approaches to studying language meaning expand and increasingly intersect, there is much to be gained from closely comparing different types and notions of meaning. Invoking Grice (1957), Campbell-Kibler (2008, 2009) observes that the role of intention seems to differ in the meanings of primary interest in (i) variationist sociolinguistics and (ii) semantics and pragmatics. The central goal of the present work is to clarify the nature of intention attribution and, in turn, the nature of these two types of meaning.

Ascribing intention. Herein, for a consequence of an action to have been INTENDED by an agent in the relevant sense means that the agent (consciously or not) performed the action as they did in part in order to effect that consequence. At the heart of this work is the question of when a consequence of an action is taken to be intended in this sense. To that question, I offer (1). Critical in (1) is the notion that actions come with various potential costs (e.g. requiring great effort to perform) and benefits (e.g. being likely to effect a desired outcomes). Depending on an agent's goals, then, a given action will appear more/less attractive.

- (1) ATTRIBUTING INTENTION. Suppose O observes agent A performing action α with potential or actual consequence c . Let M be the set of alternative actions that O believes A would have thought were available to A and more likely than α to effect c ; and let L be the set of alternative actions that O believes A would have thought were available to A and less likely than α to effect c . O is more likely to believe A intended to effect c via α :
 - a. The more O thinks *ex ante* that A would view c favorably
 - b. The more likely O thinks A would have thought α was to effect c
 - c. The less O thinks A believed A was forgoing by selecting α over elements of M
 - d. The more O thinks A believed A was forgoing by selecting α over elements of L

The first factor listed in (1) is perhaps obvious: We're more likely to think A intended to effect c the more we think that A would desire c . If we think c runs counter to A 's goals, we have relatively little reason to believe that A would try to effect c . (1b) is similarly straightforward. If, for instance, we believe A had no idea that α might effect c , we have no reason to believe that A performed α to effect c .

(1c) and (1d) center on how the action relates to alternative actions. (1c) says that our believing that A intended some outcome c is inversely related to how much we think A thought A was forgoing (in terms of benefits less costs) by selecting α over available alternatives that were apparently more likely than α to effect c . To illustrate, imagine a scenario in which there's an alternative α' that's nearly identical to α except that we think that A views α' as far more likely to effect c while being less costly. α' might then appear to have a good deal to offer relative to α , being very similar to the action A opted for and at a lower cost, and, as predicted by (1c), we would have reason to doubt that A intended for α to effect c : if effecting c were an important goal for A , why not opt for the alternative far more likely to bring that about and save in terms of costs? A similar logic underlies (1d), which says that we're more likely to think A intended to effect c the more we think A would think alternatives less likely to effect c had to offer relative to α . Consider a case with an alternative that's essentially the same as α except that we think A would believe it to be only a bit less likely than α to effect c but far less costly—making it an ostensibly attractive option. In that case, A 's opting for α offers evidence that effecting c was important to A , since A incurred far greater cost for just a small increase in the likelihood of effecting c .

The remainder of the paper examines a series of implications of (1) for different kinds of linguistic meaning, focusing especially on how the meanings of primary interest in variationist sociolinguistics differ from those of primary interest in semantics and pragmatics. A subset of the implications are presented below.

Intentionality in sociophonetic and semantic meaning. Variationist sociolinguistics has increasingly focused on how the connotations of a given form relate to the distribution of its use (e.g. Eckert 2008). The idea is that over time forms accrue social significance based on the contexts in which they appear, and, in turn, speakers wishing to evoke that social significance will have reason to employ forms that bear those associations. A classic case is the apical *-in'* form of the English *-ing* suffix, the former of which is associated with things like casualness and lack of education (Campbell-Kibler, 2008), making it potentially useful where one wishes to seem approachable, but potentially unattractive where one wishes to display, say, learnedness.

Recent work (Burnett, 2017, 2019) has attempted to capture these dynamics from the perspective of game-theoretic pragmatics. Roughly speaking, in those frameworks, the dynamics are framed as a cooperative game whereby the speaker *S* attempts to signal to the hearer *H* something *c* about *S*'s persona based on the phonetic form that *S* selects, where both *S* and *H* win if (i) *H* realizes that *S* is attempting to signal *c* and (ii) updates their beliefs to include *c*. But as (1) illuminates, that approach captures only a special case of how social significance is implied/inferred in practice. E.g., not every inference drawn from the way a person speaks is assumed to be intended. Suppose e.g. *S* is speaking a dialect of English that sounds to hearer *H* like a U.S. dialect and that, as far as *H* knows, for *S* to use a substantially different phonology would require greater effort and sound forced. *H* may then infer that *S* is from the U.S., but, barring special contexts, has little reason to believe that *S* intended to suggest U.S. nationality by their phonetics. (1) explains why. Again, as far as *H* knows, alternatives considerably less likely to suggest that *S* is from the U.S. would require more effort and sound forced. In turn, such alternatives apparently have little to offer *S*. Thus, by (1d), there's little reason to think that *S* specifically intended to signal being from the U.S. phonetically. This is a very general dynamic: believing a speaker *S* intended to effect something *c* by their utterance is facilitated by believing that *S* had an attractive alternative available that *S* thought would be less likely to effect *c*. Otherwise, *S*'s choice and form of utterance might have been motivated entirely by considerations orthogonal to the likelihood of effecting *c*. By the same token, *S* forgoing an apparently attractive alternative in favor of an utterance more likely to effect *c* provides some evidence that effecting *c* was among *S*'s goals.

Two crucial things follow (among others), *ceteris paribus*: (i) speakers are more likely to be taken to intend something by their phonetics when it appears that they're deviating from their normal phonology (cf. Podesva 2011); and (ii) meaning based in the connotations of phonetic forms is more amenable to being perceived as unintended than that derived from a form's semantic content. Speaking requires phonetics, which opens the question of whether the phonetic nature of the utterance was simply in service attempting to articulate the relevant morphosyntactic objects without expending excessive effort. But the picture is different for the semantics of morphosyntactic objects. Being convinced that an aspect of an utterance wasn't intended to suggest anything at all is facilitated by being convinced that, given the speaker's other goals and constraints, there was no reasonably good alternative to that aspect available. Morphosyntactic objects in most cases do have accessible alternatives, often including saying nothing at all. It is indeed generally very strange (with some exceptions to be discussed) to issue a morphosyntactic object, with all of its encoded semantic meaning, with no intention of suggesting something involving that semantic meaning.

By the same token, I argue that carefully considering the role of intention attribution in phonetically-based social meaning sheds light on what we read into the semantic content of others' utterances and lays bare the contingent and performative nature of all meaning in practice. For instance, it is clear that achieving one's sociophonetic ends sometimes requires having one's intention to do so go *unrecognized*. If one wishes to sound, say, cool, one had better hope that their intent to do so goes undetected—trying to seem cool is the antithesis of cool. Perhaps less obviously, however, carrying off one's intended communicative goal based in semantic content means delivering a credibly sincere performance, which, as *Hamlet's* Queen Gertrude shows us, in part means not overdoing it.

A richer understanding of the attribution of intention thus illuminates the dynamics underlying meaning in practice—from the unintended to the intended; from the sociophonetic to the semantic.

References

- Burnett, Heather. 2017. Sociolinguistic interaction and identity construction: The view from game-theoretic pragmatics. *Journal of Sociolinguistics* 21(2). 238–71.
- Burnett, Heather. 2019. Signaling games, sociolinguistic variation and the construction of style. *Linguistics and Philosophy* 42. 419–450.
- Campbell-Kibler, Kathryn. 2008. I'll be the judge of that: Diversity in social perceptions of (ING). *Language in Society* 37. 637–659.
- Campbell-Kibler, Kathryn. 2009. The nature of sociolinguistic perception. *Language Variation and Change* 21. 135–156.
- Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12(4). 453–476.
- Grice, H. Paul. 1957. Meaning. *The Philosophical Review* 66(3). 377–388.
- Podesva, Robert J. 2011. Salience and the social meaning of declarative contours: Three case studies of gay professionals. *Journal of English Linguistics* 39(3). 233–264.

Speakers' preference for more versus less-transparent causatives: Computational modeling, grammaticality judgment and production data from English, Hebrew, Hindi, Japanese, K'iche' Mayan and Balinese.

Ben Ambridge
University of Liverpool, UK
ESRC International Centre for Language and Communicative Development (LuCiD)

Laura Doherty
University of Liverpool, UK

Ramya Maitreyee
University of Liverpool, UK

Tomoko Tatsumi
Kobe University, Japan

Shira Zicherman,
Hebrew University of Jerusalem, Israel

Pedro Mateo Pedro
Universidad del Valle de Guatemala

Ayuno Kawakami
University of Liverpool, UK

Amy Bidgood
University of Salford, UK

Clifton Pye
University of Kansas

Bhuvana Narasimhan
University of Colorado, Boulder

Shira Zicherman
Inbal Arnon
Dani Bekman
Amir Efrati
Hebrew University of Jerusalem

Sindy Fabiola Can Pixabaj
Mario Marroquín PelízMargarita Julajuj Mendoza,
Universidad del Valle de Guatemala

Soumitra Samanta
University of Liverpool, UK
ESRC International Centre for Language and Communicative Development (LuCiD)

Colin Bannard
University of Liverpool, UK

Seth Campbell
University of Calgary

Stewart McCauley
University of Iowa

Ruth Berman
Tel Aviv University

Dipti Misra Sharma
Indian Institute of Information Technology, Hyderabad

Rukmini Bhaya Nair
Indian Institute of Technology, Delhi

Kumiko Fukumura
University of Strathclyde

This work investigates the ways in which six different natural languages use morphosyntactic marking to mark difference degrees of agency, and the continuum from intentional and/or directly-caused to accidental and/or indirectly-caused events. Many of the world's languages (Shibatani & Pardeshi, 2002, discuss 38 examples) have (at least) two causative structures, which particular verbs prefer to a greater or lesser degree: (1) a *more-transparent* structure with the verb *cause/make/do* or a morpheme that is often a historically grammaticalized form of that verb (e.g., Japanese *-(s)ase*), and (2) a *less-transparent* structure that marks causation more idiosyncratically (e.g., by using a form that is either indistinguishable from a non-causative or stem form, or similar to such a form, but with a vowel or consonant change that is only partially predictable):

- (1) Someone made the truck break [More transparently-marked causation]
- (2) Someone broke the truck [Less transparently-marked causation]

The formal realizations of these different forms vary from language to language. English, as illustrated by the examples above, relies primarily on syntax. Japanese, Hindi and K'iche rely primarily on morphology, in the form of a more-transparent causative marker (*-(s)ase*, *-aa*, and *-(i)sa-j* respectively) and various types of less-transparent stem-change. For Hebrew, the root is defined as a three-consonant (C.C.C) pattern (e.g., *sh.b/v.r* for BREAK), which forms a verb only when it is inserted into a binyan template; in this case either the dedicated causal binyan *hiCCiC* (e.g., *hishbir*) or the appropriate general transitive binyan: CaCaC (e.g., *shavar*) CiCeC or hiCaCeC, for the more- and less-transparent forms respectively.

The aim of this study was to investigate how speakers learn which verbs prefer which causative form (more-/less-transparent), and to what extent. First, 20 adult native speakers of each language rated each of 60 actions for four semantic properties relative to the notion of agency (from Shibatani & Pardeshi, 2002):

Event-merge: The extent to which the causing and caused event are two separate events or merge into a single event that happens at a single time and a single point in space

Autonomy of the causee

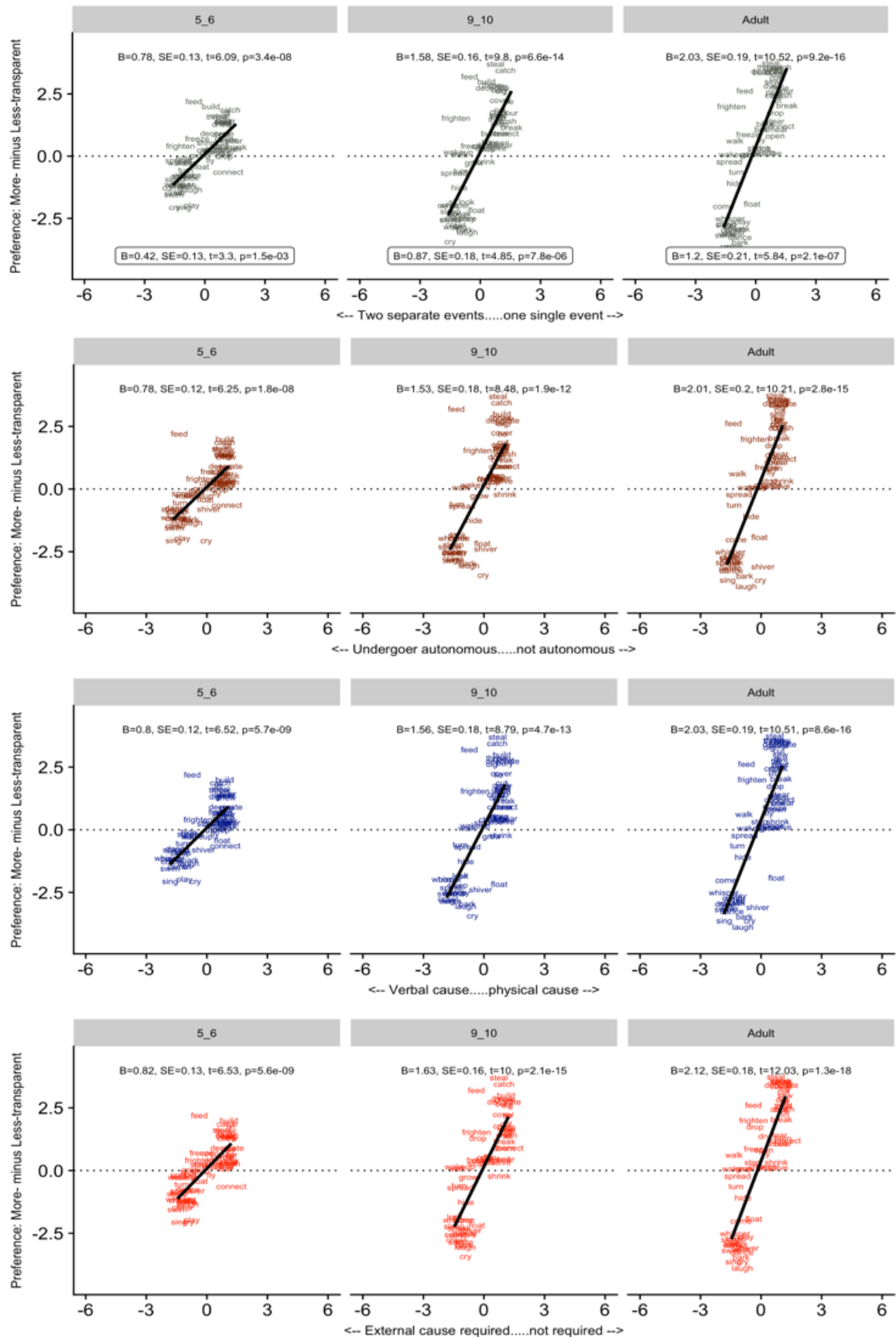
Requires: Whether the caused event requires a causer

Directive: Whether causation is directive (e.g., giving an order) or physical

These ratings, together with corpus frequency information, were to build a computational model of speakers' causative preferences for each language. For every language except K'iche', this model significantly predicted speakers' preference for the more- versus less-direct causative form across 60 verbs, as assessed using continuous grammaticality judgments (adults, children aged 5-6 and 9-10), binary grammaticality judgments (4-5) and elicited production (4-5 and 5-6) ($N=48$ participants per age group per study). The model was also able to generalize to unseen verbs on the basis of their semantics (when a randomly-selected half of the verbs were presented to the model for the first time at test), and even to ratings from an entirely new language, on which the model had not been trained (Balinese). We conclude that the morphosyntactic realization of causativity is determined primarily by the degree of conceptual merging between the causing and caused event (Shibatani & Pardeshi, 2002), and that children have some awareness of this fact by at least 4-5 years of age.

Shibatani, M. & Pardeshi, P. (2002). The causative continuum. In *The Grammar of Causation and Interpersonal Manipulation*, Masayoshi Shibatani (ed.), pp. 85–126. Amsterdam: John Benjamins.

Figure 1. Example data from English. Participants' preference for the more-transparent (periphrastic *make* causative) over the less-transparent (transitive causative) form for each verb, as a function of the semantic predictors of (a) Event-Merge, (b) Autonomy of the causee, (c) Whether causation is directive or physical and (d) whether the caused event requires a causer.



Note: The values shown for each predictor are from nonpartial (single predictor) models only

Agency and agents in the English WAY-construction

Caterina Cacioli
University of Florence

The WAY-construction (exemplified in sentence 1) generally describes a movement along a path. Specifically, in the literature (Goldberg 1995) it has been constrained to a self-propelled movement:

(1) I made my way through the crowd

Being the action self-propelled, there must be an agent capable of engaging in a movement. Interestingly, self-propelled motion is not constrained to intentional movements or movements accomplished by an animate agent: the action need not be volitional and inanimate entities participate in the construction as long as motion is *construed* as self-propelled (Goldberg 1995):

(2) The seeds pushed their way through crusted soil

(3) *The wood burned its way to the ground

(Goldberg 1995)

Moreover, the construction has been constrained to unergative verbs (Goldberg 1995; Levin and Rappaport 1995) thus ruling out the unaccusative ones, which are associated with lack of agentivity and self-initiation (Van Valin 2001).

A dedicated research on the COCA (Corpus of Contemporary American English) allowed to gather a large quantity of data to point up the relation between intentionality of action, self-propelled movement and inanimate entities. Acceptability judgement tests were also used to furtherly illustrate this relation.

Following Oyon (2013) I argue and show with data that motion involved in the construction does not need to be self-propelled – simply, the subject must be able to undergo or perpetuate motion. Moreover, it is shown that while unergative verbs are highly preferred, unaccusative verbs are also used¹, as for example the verb *grow*. In particular, they are used when they denote a specific manner (as with *shrink* or *roll*). In explaining the results from acceptability judgements, it may be argued that the acceptability of the sentence increases when the action is construed as intentional:

(4) Mary wasn't paying attention and fell her way down the stair.

(5) Mary wasn't paying attention and tripped her way down the stair.

(6) The skater wanted to make a new record but tumbled his way down a high ramp.

¹ Narasimhan (2003) discusses the ability of English manner of motion verbs to be construed as unergative or unaccusative.

These sentences scored a 5-point Likert scale mean of, respectively, 1.91, 2.96 and 3.27. *Fall* is not attested and neither acceptable, while *trip* and *tumble* are attested and seem to receive higher acceptability judgments based on the intentionality of the action in which they are used: in (5) the subject is not actively performing an action, while the subject in (6) - the skater - is.

Intentionality of action is also most importantly at stake when the subject is inanimate. Occurrences with inanimate entities are construed around a context of *perceived intentionality*. Certainly, the actions in (7-9) are not intentional, but are perceived as such:

- (7) The doughnut quickly worked its way into American culture (COCA, San Fran Chron 2004)
- (8) [the magma] is going to melt its way to the surface and to destroy us [...] (NOW, The daily Galax, 2013)
- (9) the group examined how a robot powered by a nuclear reactor would melt its way through the ice (NOW, yahoo news 2018)

Conversely, the subject in sentence (3) is not perceived as intentional: the wood is moving in a downward direction while burning because of an evident causer, the fire. Instead, sentences (7-9) may be perceived as intentional in lack of an overt causer of the action. I argue that the variation in the acceptability of sentences (7-9) and (3) relies on humans' capability of conceptualizing actions and attributing intentionality.

References

- Goldberg, A. (1995). *A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.
- Israel, M. (1996). The Way Constructions Grow. In A. Goldberg, *Conceptual Structure, Discourse and Language*, (pp. 217-230). Stanford: CSLI Publications.
- Levin, B., & Rappaport, M. (1995). *Unaccusativity at the Syntax-lexical Semantic Interface*. Oxford: The Mit Press.
- Narasimhan, B. (2003). Motion events and the lexicon: a case study of Hindi. *Lingua*, 113, 123-160.
- Oyón, A. (2013). Revisiting Goldberg's semantic constraints on the 'way' construction. *Revista Española de Lingüística Aplicada*, 26, 349-364.
- Perek, F. (2018). Recent change in the productivity and schematicity of the way-construction: A distributional semantic analysis. *Corpus Linguistics and Linguistic Theory*, 14, 65-97.
- Van Valin, R. (2001). *An Introduction to Syntax*. Cambridge: Cambridge University Press.

Skill, luck, and mere actions in grammar

Bridget Copley, SFL (CNRS/Paris 8)

Cova (2016) proposes what we might call an “intentional pluralism” account of intentional action, where “intentionally” can have the following construals: (A) A positive meaning, according to which someone does something intentionally when he actively does it based on his or her desire to do it; (B) A ... meaning according to which someone does something intentionally when he does it without being forced to do it. In this sense, “intentionally” is opposed to “unwillingly” or “by force”; (C) A ... meaning according to which someone does something intentionally when he does it by having full control upon his action. In this sense, “intentionally” is opposed to “by accident” or “by sheer luck”. In this talk I will consolidate (if not quite unify) these three meanings, and align them very broadly with cross-linguistic facts in the Austronesian, South Slavic, and Salish language families. To achieve these goals I present an account of how action is represented in natural language grammar, relying on particular treatments of causation and intention. We end up with several ways to refer to action: mere action on the one hand, and on the other, four kinds of dispositionally-caused action based on the features internal/external and intention/disposition. In “skill” cases, the cause is an intention internal to the actor; in “luck” cases the cause is a disposition external to the actor.

Mere actions: It is possible in English and many other languages to refer to a “mere” action, that is, without regard to whether it is intentional or accidental. This gives us a sense of agenthood that is also independent of intentionality or accidentality. “Process” theories of causation that give us access to representations of energy transfer (e.g. Kistler 2006, Wolff 2007) can be useful here, in that an agent might be understood as the source of the energy that goes into the action (Copley & Harley 2015). (The notion of agent that is needed will be further elaborated.) Mere actions can themselves be directly caused, either by another action, or by a state such as the intentional state of the agent. The directness of the causation is linguistically significant; see Martin 2018.

Dispositionally-caused action: Intentions, I argue, are a kind of disposition. That is, intentions share with “ordinary” dispositions a formal structure even though they have interpretational differences. The evidence for this lies in the fact that intentions and dispositions are often treated the same by grammar (Copley 2018). For instance, *have*-causatives in English can take animate subjects or inanimate subjects, as shown in (1). If an inanimate subject, it is a property/state/causal power of the subject that causes my laughter. If an animate subject, it can similarly be a property of the subject as in (1a), or it can be an intentional state of the subject, as in (1a), where Mary has authority to impose her intention on my action. Intentions can cause more things than ordinary dispositions, as we can see from the fact that (1b) is not possible with an inanimate subject.

- (1) a. Mary/the book had me calling my mother.
- b. Mary/*the book had me call my mother.

Such cases have been discussed in the literature for Vendlerian accomplishments (Folli & Harley 2005) and nominalizations (Sichel 2010; Alexiadou et al. 2013), among others. Intentions and ordinary dispositions can thus be argued to share a formal structure: a state s is (i) held by an individual x , and is (ii) intensionally directed toward a (propositional) goal p , and (iii) causes an event e (under certain circumstances C , as per Fara 2001, but these will not be important in the cases we are looking at). Grammars may, or may not, distinguish intentions from other dispositions. Intention and ordinary disposition are *not* the two kinds of dispositional causes for mere actions that I have promised, however. Instead it is whether the causing dispositional state is internal or external to the actor. If the causing state is internal, it can be either an intention or an “ordinary” disposition; either one I call “Skill”; so the cases in (1) are to be considered cases of Skill. If the causing state is external and an ordinary disposition, we are dealing with “Luck”.

Consolidating Cova’s proposal: This picture of mere action, Skill, and Luck, if correct, would allow us to consolidate Cova’s three meanings for “intentionally”. (A) refers to action directly caused by an (intentional) disposition internal to the actor, i.e. Skill. (B) refers to a mere action which is directly caused by a mere action or intention external to the actor. (C) refers to an action not caused by any external disposition, i.e., not Luck. So (A) specifies that the direct cause for the action is internal (and an intention), while (B) and (C) specify that the direct cause for the action is not external. (A) and (B)+(C) may be equivalent in the ontology of English functional morphology, but perhaps not cross-linguistically, because of the possibility for internal ordinary dispositional causes and external intentional causes as well as totally uncaused mere actions.

Cross-linguistic evidence Languages can grammatically distinguish internal/external and intentional/dispositional causes in their episodic verb forms. (Additional distinctions may also be reflected.) External intentions and internal dispositions (e.g. Skwxwú7mesh out of control forms (Jacobs, 2007) and South Slavic involuntary state constructions (Rivero 2009)) are also possible causes. Luck forms are, like English lexical “manage to”, compatible with both bad luck and good luck, and mere action forms, when competing with dispositional forms, have a “suddenly” use where the action is understood to be uncaused. I argue that Tagalog distinguishes between a Neutral verb form that gets a Skill reading, and an Ability/Involuntary action form that gets Luck readings ((2)). The St’át’cimets out of control form, I propose, conveys either mere action (“suddenly”) or Luck ((3)).

(2) Skill vs. Luck: Tagalog (Alonso-Ovalle & Hsieh 2017)

- a. B[**in**]uks-an ni Lisa ang pinto.
[PFV.NTL] open-LV GEN Lisa NOM door
‘Lisa opened the door (deliberately).’
- b. Na-**buks**-an ni Lisa ang pinto.
PFV.**AIA**-open-LV GEN Lisa NOM door
‘Lisa managed to open the door. / Lisa accidentally opened the door.’

(3) One form for mere action and Luck: St’át’cimets out of control form (Davis et al. 2009)

- a. **ka**-lhexw-a ta=n-sqax7=a
ka-appear-**a** det=1sg.poss-dog=exis
‘My dog appeared suddenly.’
- b. **ka**-qám’t-s=kan-**a** ta=wa7 ts’aq’-n-an
ka-hit-caus=1sg.subj-**a** det=impf throw-dir-1sg.erg snimulh1 1pl. emph.
‘I managed to hit the target.’

Both contrast with English, in which episodic verbs typically express mere action (with some lexical inclusion of intention in verbs such as *murder*). Intriguingly, the acquisition of the implicit understanding that the English verb forms are mere action forms, and not Skill forms, in the absence of contrasting grammatical forms may not be straightforward, as shown in (4):

(4) Zoe (3;10), replying to “Someone lost their hat”: “Why did they want to lose their hat?”

Selected references: Alonso-Ovalle, L. & H. Hsieh 2017. Causes and expectations: On the interpretation of the Tagalog ability/involuntary action form. SALT 27. Copley, B. 2018. *Dispositional causation. Glossa*, 3(1). Cova, F. (2016 ms.). The folk concept of intentional action: Empirical approaches. In W. Buckwalter J. Sytsma (Eds.), *Blackwell Companion to Experimental Philosophy*. . Davis, H. et al. 2009. Out of control marking as circumstantial modality in St’át’cimets. In L. Hogeweg, H. de Hoop, and A. Malchukov (eds.), *Cross-linguistic semantics of tense, aspect and modality*.

Modal vs. Non-modal ‘Un-agentive’ constructions in Laz

Ömer Demirok, Boğaziçi University

Typological and theoretical background. The understood agent in Involuntary Agent Constructions (IAC), semantically, appears to lack control over the event, and morpho-syntactically, bears non-canonical marking, which may also be indexed on the verb somehow (Kittilä 2005). Importantly, IAC exhibits polysemy cross-linguistically. Kittilä notes that IAC may, in some languages, also describe situations where an agent acts intentionally yet has ‘a lower degree of control’, citing situations where the agent ‘finally or unexpectedly manages to do something’.

Broadly speaking, two theoretical perspectives on the syntax and semantics of IAC have emerged:

(i) Schäfer (2009) proposes an analysis for IAC (in e.g. German, Greek, Agul) where the event is an anticausative change-of-state event and there is an applicative projection on top of it, introducing a **possessor** rather than a unintentional causer. Schäfer argues that the vague semantics of possession will accommodate both ‘unintentionally do X’ and ‘manage-to do X’ readings.

(ii) Davis et al. (2009) propose that the construction that got cited as an example of IAC in St’át’imcets feature a **circumstantial modal** which has the ability reading when its force is existential and the compulsion reading when its force is universal.

Then, the obvious question: *do we need both of these non-modal and modal analyses?*

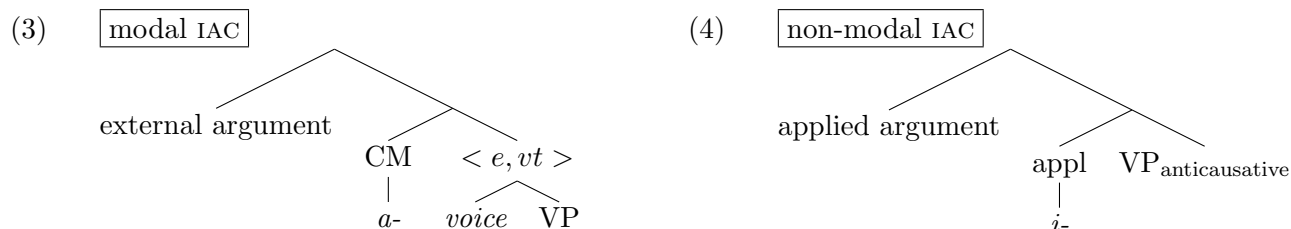
This paper argues that two distinct IAC-like constructions co-exist in Laz, suggesting both lines of analyses are independently needed.

Data. Laz¹ (endangered, South Caucasian, spoken in Turkey) has two IAC-like constructions, which I will call modal IAC and non-modal IAC, respectively (on the former, see Öztürk (2013), Demirok (2018)). The apparent morphosyntactic contrast is located in the pre-root prefix. I argue that *a-* in (1), a modal IAC, expones a circumstantial modal whereas *i-* in (2), a non-modal IAC, expones an applicative head.

(1) Ma mak’vande-s para me- m- a- ç -u
 1SG.DAT beggar-to money.NOM PV- 1SG- CM- give -PST
 ‘I couldn’t but give money to the beggar.’ ⇐ modal IAC

(2) Ma mtsxuli me- m- i- l -u
 1SG.DAT pear.NOM PV- 1SG- APPL- GO -PST
 ‘I (accidentally) dropped the pear.’ ⇐ non-modal IAC

In particular, I propose that modal IAC and non-modal IAC have the following structures in Laz. The former features a circumstantial modal that combines with an unsaturated voiceP, whereas the latter features an applicative head that combines with an unaccusative change-of-state event.



There are a number of distributional observations that justify this asymmetry in the event composition — setting aside modality for now. Three of them: (i) modal IAC can embed unergatives (which feature *voice*) (ii) modal IAC cannot embed unaccusative events (which do not contain *voice*) (iii) the regular causativizer is required for modal IAC to embed a lexically unaccusative verb.

For reasons of space, here I only provide comparisons with baseline structures below.

¹Laz data drawn from author’s own current and previous fieldwork notes.

- (5) baseline → modal IAC
- a. Himu-**k** mtsxuli tsad-u
3SG-ERG pear look-PST
'She looked at the pear.'
- b. Himu-s mtsxuli **a**-tsad-u
3SG-DAT pear CM-look-PST
'She couldn't *not* look at the pear.'
- (6) baseline → non-modal IAC
- a. Ditsxiri do-kort-u
blood PV-clot_{intr}-PST
'The blood clotted.'
- b. Ditsxiri do-m-**i**-kort-u
blood PV-1SG-APPL-clot_{intr}-PST
'I (accidentally) let the blood clot.'

Why need both modal and non-modal?

First, Modal IAC in Laz *requires* an (unsaturated) voiceP, which entails that the DAT NP it features is semantically the external argument (under a (trivial) semantic composition reminiscent of the PASS head in Bruening 2012). A supporting point for this is that instruments are licensed as the DAT argument in modal IAC - shown in (7). This is not possible in non-modal IAC. [not shown here]

- (7) a. Xami-k xe-šk'imi k'vat-u
knife-ERG hand-my cut_{trns}-PST
'The knife cut my hand'
- b. Xami-s xe-šk'imi **a**-k'vat-u
knife-DAT hand-my CM-cut_{trns}-PST
'The knife couldn't *not* cut my hand'

While it is not surprising for a root-modal to be agentive (Mandelkern et al. 2017), this is highly surprising if we wanted to subsume modal IAC under the well-studied form of IAC which *requires* an anticausative event, as Fauconnier (2011) and Schäfer (2009) both stress.

Second, modal IAC exhibits polysemy, as Kittilä argues IAC in general does. Yet, on a closer look, the polysemy in modal IAC is fully within the confines of a circumstantial modal, allowing readings where an event is allowed or forced to unfold by its circumstances. Nothing beyond that.

Notably, modal IAC in Laz is force-variable, as shown in (8b), like its counterpart in St'át'imcets. In addition, there is evidence that CM is apparently an existential modal, as its behavior under negation suggests. Note that (8c) is not ambiguous.

[Though, it is not clear if its universal construal is to be derived via strengthening, or is only available due to entailment (see Deal (2011) for Nez Perce).]

- (8) a. Bere-k opša šk'om-u
child-ERG a lot eat-PST
'The child ate a lot.'
- b. Bere-s opša **a**-šk'om-u
child-DAT a lot CM-eat-PST
i. 'The child was able to eat a lot.'
ii. 'The child couldn't *not* eat a lot.'
- c. Bere-s opša var **a**-šk'om-u
child-DAT a lot NEG CM-eat-PST
✓'The child wasn't able to eat a lot.'
✗'It isn't the case that
the child couldn't *not* eat a lot.'

Third, there is evidence that the syntax of the VP domain isn't the only point of difference between non-modal IAC and modal IAC. If this were the case, we would expect them to have identical construals in all environments where they are talking about the same event. Here is how I test this: Laz has distinct roots for 'break', one lexically transitive, one lexically unaccusative. As expected, they go with modal IAC (9a) and non-modal IAC (9b), respectively. Yet, they don't converge on the same meaning, it seems. Under negation, (9a) only gives us a construal that we expect from a possibility modal, which (9b) lacks.

- (9) a. Ham kva va m-**a**-t'ax-u
this stone NEG 1-CM-break_{trns}-PST
'I wasn't able to break this stone.'
- b. Ham kva va m-**i**-t'rox-u
this stone NEG 1-APPL-break_{intr}-PST
≈ 'I didn't break this stone.'

Selected References:

- Davis, H, L. Matthewson, & H. Rullmann. 2009. Out of control marking as circumstantial modality in St'át'imcets. In *Cross-linguistic Semantics of Tense, Aspect and Modality*.
- Seppo Kittilä (2005) Remarks on involuntary agent constructions, *Word*.
- Öztürk, B. 2013. Low, high and higher applicatives. In *Linguistik Aktuell/Linguistics Today 197*.
- Schäfer, F. 2009. The Oblique Causer Construction across Languages. In *Proceedings of NELS 38*.

Intentions and the Future in Buridan's Bridge

Paul Egré (Institut Jean-Nicod, PSL University)

Abstract: A chapter of Cervantes's *Don Quixote* (II, 51) confronts Sancho Panza to a version of the Liar Paradox, also known as Buridan's Bridge (*Sophismata*, chap. 8, sophism 17). The law declares that whoever comes to cross a certain bridge will be hanged at the gallows if they make a false statement as to the goal of their passing, and will pass free if they tell the truth. Then, "one man said that by the oath he took he was going to die upon that gallows that stood there, and for nothing else" ("iba a morir en aquella horca que allí estaba, y no a otra cosa"). The judges face a quandary: if the man passes free, he will have lied, and must be hanged; but if he is hanged, he will have spoken truly, and should be set free. The goal of this paper is twofold: my main motivation is to examine the sense in which the problematic sentence can be considered both true and false, in line with dialetheism (Priest 2006), and with the strict-tolerant account of the Liar (Ripley 2012, Cobreros et al. 2013), but also with earlier remarks by Jacquette (1991), suggestive of a dialetheist analysis. A subordinate goal, more specifically related to the topic of the conference, is to assess whether taking into consideration the utterer's intention is of any help in solving the paradox. Buridan's own solution is that because it pertains to a future contingent, the utterance is neither true nor false (see also Ulatowski 2003). In Cervantes' case, Sancho judges that the utterance is both true and false, and lets the man pass freely. One issue there is whether this outcome is compatible with the man's oath and intention to die at the gallows, assuming the utterance is true. I will discuss whether and how a dialetheist account can deal with this problem.

On abstinence and avoidance

Maria Esipova (University of Oslo)

Intro I inspect two types of readings emerging, e.g., for negated imperatives, ‘not want’, and ‘fear’: (i) ABSTAIN readings, implying willingness to prevent someone from intentionally engaging in an activity, and (ii) AVOID readings, implying willingness to prevent an unintended outcome:

- (1) a. Don’t call anyone! / I don’t want to call anyone. / I’m afraid of calling anyone. ABSTAIN
 b. *In a potential butt-dialing context:* Don’t (accidentally) call someone! / I don’t want to (accidentally) call someone. / I’m afraid of (accidentally) calling someone. AVOID

In English, the distinction manifests in, e.g., anti-/re-licensing of *some* (observed in Szabolcsi 2004): (1). I focus on aspect and indefinites in ABSTAIN vs. AVOID cases in Russian, but the proposed approach extends to English.

Core puzzles (2)–(4): 1. Across the three environments, verbs are in imperfective in ABSTAIN cases and in perfective in AVOID cases. 2. Under negated imperatives and ‘not want’, only *ni*-NPIs are licensed in ABSTAIN cases, but both *ni*-NPIs and dependent *nibud’* indefinites (discussed, e.g., in Pereltsvaig 2008) are licensed in AVOID cases. Under ‘fear’, only free choice indefinites (e.g., *libo*) are licensed in ABSTAIN cases, but both free choice and *nibud’* indefinites are licensed in AVOID cases.

- (2) a. Ne {zvoni / *pozvoni} {nikomu / *komu-nibud’}!
 not {call.IPFV.IMP / *call.PFV.IMP} {NI-who / *who-NIBUD’}
 ‘Don’t call anyone!’ ABSTAIN
 b. Ostorozhno! Ne {*zvoni / pozvoni} (slučajno) {nikomu / komu-nibud’}!
 careful not {*call.IPFV.IMP / call.PFV.IMP} (accidentally) {NI-who / who-NIBUD’}
 ‘(Careful!) Don’t (accidentally) call someone!’ AVOID
- (3) a. Ja ne xoču {nikomu / *komu-nibud’} {zvomit’ / *pozvomit’}.
 I not want {NI-who / *who-NIBUD’} {call.IPFV.INF / *call.PFV.INF}
 ‘I don’t want to call anyone.’ ABSTAIN
 b. Ja ne xoču {nikomu / komu-nibud’} (slučajno) {*zvomit’ / pozvomit’}.
 I not want {NI-who / who-NIBUD’} (accidentally) {*call.IPFV.INF / call.PFV.INF}
 ‘I don’t want to (accidentally) call someone.’ AVOID
- (4) a. Ja bojus’ {komu-libo / *komu-nibud’} {zvomit’ / *pozvomit’}.
 I fear {who-LIBO / *who-NIBUD’} {call.IPFV.INF / *call.PFV.INF}
 ‘I am afraid of calling anyone.’ ABSTAIN
 b. Ja bojus’ {komu-libo / komu-nibud’} (slučajno) {*zvomit’ / pozvomit’}.
 I fear {who-LIBO / who-NIBUD’} (accidentally) {*call.IPFV.INF / call.PFV.INF}
 ‘I’m afraid of (accidentally) calling someone.’ AVOID

Prior work Aspect in negated imperatives in Slavic was recently analyzed in Goncharov 2018, which relied on an ad hoc intentionality operator in ABSTAIN cases. English *some* under ‘not want’ was independently analyzed in a more pragmatic way in Goncharov 2020. While I share some of Goncharov’s insights, I propose a more principled and general analysis, with both the aspect and the indefinite facts following from distinct compositional structures in ABSTAIN vs. AVOID cases, and pragmatics affecting the choice between the two.

Proposal Events vs. situations I distinguish *events* from *situations*. Situations can be thought of as (potentially) complex events whose parts are situated within the situation’s runtime a.k.a. the reference time (RT_s). For concreteness, I assume the following mereology: an event can be part of a situation ($e \subseteq s$) or a world, and a situation can be part of a world. \exists binding the event variable always takes the lowest scope, which I assure by adopting (ontologically enriched) Champollion’s (2015) quantificational events semantics, where verbs denote existential quantifiers over events (e.g., $[[\text{‘open’}]] = \lambda f_{vt} \lambda w_s. \exists e_v. f(e) \wedge e \subseteq w \wedge \text{open}(e)$), and any further modification of e is done via the continuation f , eventually closed off via a trivial continuation, $\lambda e.T$. In contrast, the situation layer is optionally introduced later in the derivation.

Situations & aspect Russian imperfective verbs can describe an event that is part of a situation, in which case the endpoints of the event’s runtime (including the projected culmination of wannabe accomplishments like ‘open.IPFV closet’) are placed outside RT_s ; but they can also describe events that are not situated relative to any RT_s . Russian perfective verbs always describe events that are part of a situation and culminate within RT_s .

Cases at hand AVOID cases involve situation descriptions, but ABSTAIN cases don't. An independent piece of evidence is that only AVOID cases allow 'once'/'one day', which I take to be a situation(-introducing) modifier:

- (5) a. Ne {*otkryvaj / otkroj} odnaždy škaf!
 not {*open.IPFV.IMP / open.PFV.IMP} once closet
 'Don't once open the closet!' (only AVOID)
- b. Ja {ne xoču / bojus'} odnaždy {*otkryvat' / otkryt'} škaf.
 I {not want / fear} once {*open.IPFV.INF / open.PFV.INF} closet
 'I don't want to / I am afraid that I will} once open the closet.' (only AVOID)

For concreteness, I assume that the imperative operator and 'want' compose with their complements in the same way in AVOID and ABSTAIN cases (this is not crucial): (6). As for 'fear', in ABSTAIN cases, it simply encodes a relation between its subject and its complement (propositional or not) in the world of evaluation, while in AVOID cases, it asserts the epistemic possibility of its propositional complement and presupposes the speaker's "fearful" attitude towards it (I think this at-issue vs. not-at-issue distinction is pragmatic and don't encode it lexically): (7).

- (6) [['Don't open the closet!' / 'I don't want to open the closet'] = 1 in w iff $\forall w' R_{imp/want} w$:
- a. ABSTAIN: $\neg[\exists e.e \subseteq w' \wedge \text{open}(e) \wedge \text{ag}(e) = \text{addr/sp} \wedge \text{th}(e) = \iota x.\text{closet}(x)]$
 b. AVOID: $\neg[\exists s.s \subseteq w' \wedge \exists t.t \subseteq \text{RT}_s \wedge \exists e.e \subseteq s \wedge \text{culm}(e, t) \wedge e \subseteq w' \wedge \text{open}(e) \wedge (\dots)]$
- (7) [['I am afraid (a.) of opening the closet / (b.) that I will open the closet'] = 1 in w iff
- a. ABSTAIN: $\text{fear}(\text{sp}, \lambda w'. \exists e.e \subseteq w' \wedge \text{open}(e) \wedge \text{ag}(e) = \text{sp} \wedge \text{th}(e) = \iota x.\text{closet}(x), w)$
 b. AVOID: $\exists w' R_{bel} w [\exists s.s \subseteq w' \wedge \exists t.t \subseteq \text{RT}_s \wedge \exists e.e \subseteq s \wedge \text{culm}(e, t) \wedge e \subseteq w' \wedge \text{open}(e) \wedge \text{ag}(e) = \text{sp/addr} \wedge \text{th}(e) = \iota x.\text{closet}(x)] \wedge \text{fear}(\text{sp}, \lambda w''. \exists s.(\dots), w)$

Indefinite licensing I assume that (i) quantification over worlds in imperatives and 'want' licenses *nibud'*; (ii) *ni*-NPIs are licensed by negation and, when licensed, block *nibud'*; (iii) an intervening existential quantification over situations blocks *ni*-NPI licensing and re-licenses *nibud'*. Thus, in the ABSTAIN cases at hand involving negation, only *ni*-NPIs are licensed, but in the AVOID cases, either one can be licensed, depending on whether the indefinite scopes above or below the situation (*ni*: $\neg[\exists x \exists e]$ or $\neg[\exists x \exists s \exists e]$; *nibud'*: $\neg[\exists s \exists x \exists e]$). ABSTAIN-type 'fear' doesn't on its own license *nibud'* indefinites (nothing for them to depend on), but AVOID-type 'fear' does, thanks to the quantification over worlds.

When to have situations In all the three environments at hand, the speaker considers certain scenarios undesirable. In ABSTAIN cases, this attitude can be about a certain outcome (e.g., they don't want anyone to get called) or the activity (potentially) leading to it (e.g., they don't want the agent to engage in any calling, as they think cell phones are harmful). AVOID cases are unambiguously about the outcome. If the speaker wants to prevent a certain outcome and thinks it can be prevented by the agent not engaging in the activity of the relevant kind, they go for an unsituated event description (ABSTAIN). But if they don't think all such events can be prevented (e.g., the relevant activity can't be controlled or might not be recognized by the agent as leading to a certain end-state), they have to go for a situation description containing the end-state (AVOID). An argument for such a pragmatics-heavy story comes from complex situation descriptions. E.g., in (8), the part of the undesirable situation outside the agent's control is whether they see a monster or hear a scraping sound; the closet opening is intentional (as indicated by the purposive clause), and the aspect choice is determined solely by its placement relative to RT_s .

- (8) a. Ja ne xoču odnaždy otkryt' škaf, čtoby dostat' noski, i uvidet' tam monstra.
 I not want once open.PFV.INF closet to get socks and see.PFV.INF there monster
 'I don't want to once open the closet to get socks and see a monster there.'
- b. Ja ne xoču odnaždy otkryvat' škaf, čtoby dostat' noski, i uslyšat' skrežet.
 I not want once open.IPFV.INF closet to get socks and hear.PFV.INF scrape
 'I don't want to once be opening the closet to get socks and hear a scraping sound.'

Expanding the discussion I also show that subject obviation in subjunctive (or indicative) clauses (Farkas 1992 et seq.) is distinct from the ABSTAIN vs. AVOID distinction, as the latter still holds when there is no obviation. In addition, I briefly discuss how the proposed approach extends to situations and lack thereof under some other predicates (e.g., 'hope') and root modals, and the connection between minimizers and aspect in Russian.

References Champollion. 2015. *L&P* 38(1). Farkas. 1992. In *Lexical Matters*. Goncharov. 2018. In *FASL* 27. Goncharov. 2020. In *SALT* 30. Pereltsvaig. 2008. In *WCCFL* 27. Szabolcsi. 2004. *NLLT* 22(2).

The ‘Knobe effect’ as an instance of a ‘severity effect’

In the last two decades, several empirical studies have documented an asymmetry in people’s assessments of intentional action, so-called ‘Knobe effect’ (cf. Knobe, 2003; Cushman & Mele, 2008; Knobe & Mendlow, 2004; Knobe & Burra, 2006; Nadelhoffer, 2004, 2005; for detailed reviews see Feltz, 2007; and Nichols & Ulatowski, 2007). Accordingly, foreseen (yet undesired) outcomes that are harmful are judged intentional, whereas foreseen (yet undesired) outcomes that are helpful are judged nonintentional. In a more recent study, however, Kneer & Bourgeois-Gironde (2017) found that people’s ascriptions of intentionality are susceptible not only to the Knobe effect but to a ‘severity effect’: the more harmful a foreseen (yet undesired) outcome, the more inclined people are to say that it was intentionally brought about.

The severity effect can be perceived as a challenge to some explanations of the Knobe effect. The latter has standardly focused but on two data points: intentionality ascriptions for harmful v. helpful (or neutral) outcomes. Not surprisingly, many scholars have conceived of the Knobe effect as a *binary*, absolute effect: foreseen harmful outcomes are deemed intentional, whereas foreseen helpful—or at least not harmful—outcomes are deemed nonintentional. However, the severity effect findings suggest that things might be somewhat more complex. Rather than being of binary nature, the relation between an action’s outcome and intentionality ascriptions appears to be a matter of *degrees*. It is not just that people judge harmful outcomes as intentional; it is that people’s propensity to say that an outcome is intentional is commensurate with the outcome’s degree of harm.

The research conducted by Kneer & Bourgeois-Gironde (2017) on the severity effect, however, suffers from a shortcoming. Focus was placed only on intentionality ascriptions for graded *harmful* outcomes (somewhat bad v. very bad); beneficial outcomes were not tested. While there is empirical evidence concerning the relation between graded harmful outcomes and intentionality ascriptions (cf. Kneer & Bourgeois-Gironde, 2017; Prochownik, 2020), the relation between differently graded positive outcomes and intentionality attributions is completely unexplored.

To address the former lacuna and provide a clearer understanding of the relation between outcomes and intentionality ascriptions, we conducted a study exploring attributions of intentionality (and knowledge) across a range of different outcomes: *very bad, somewhat bad, neutral, somewhat good and very good*. In contrast to classic side-effect effect research, we were less interested in the difference between good and bad outcomes, but in the correlation between perceived goodness and badness of outcome (which we also measured) on intentionality (and knowledge). Given that we contend that perceived blameworthiness (or praiseworthiness), rather than outcomes per se, might drive intentionality ascriptions, we were also interested in the correlation between perceived blame and praise on intentionality. Importantly, we decided to explore the positive and the negative part of the outcome spectrum, as well as the blame and praise parts, separately. In the next paragraph we will briefly summarize the results of Experiment 1 of our study.

On the negative part of the outcome spectrum, consistent with the severity-effect findings by Kneer & Bourgeois-Gironde (2017), we found a positive correlation between intentionality ascriptions and perceived outcome badness ($r=.416, p < .001$): the worse participants perceived the side effect to be, the more likely they were to agree with the claim that the side effect was

intentionally brought about. On the positive side of the outcome-spectrum, by contrast, we found a weak negative correlation between intentionality ascriptions and perceived outcome goodness ($r = -.192, p = .013$): the more desirable participants perceived the side effect to be, the less likely they were to agree with the claim that it was intentionally brought about. Interestingly, we also found a strong positive correlation between blame and intentionality ascriptions ($r = .741, p < .001$), and a weaker positive correlation between praise and intentionality ascriptions ($r = .385, p > .001$).

The former results suggest that the Knobe effect data points are but two data points of a broader, more fine-grained phenomenon, and that Knobe effect explanations that have conceived of it in binary terms are at best incomplete. Explanations that, on the contrary, turn to gradable features such as blame or praise, seem to be on the right track. The results further suggest that the relation between intentionality ascriptions and graded outcomes is different when the valence of the latter is negative than when it is positive. Whereas an increase in the degree of harm of a foreseen outcome warrants a considerable increase in the propensity to ascribe intentionality to it, an increase in the degree of desirability of a foreseen outcome only warrants a minor (if not irrelevant) increase in people's propensity to judge it nonintentional. The results of Experiment 2, which will also be presented, demonstrate that these findings are robust.

References:

- Cushman, F., & Mele, A. (2008). Intentional action: Two and a half folk concepts? In J. Knobe, & S. Nichols (Eds.). *Experimental philosophy* (pp. 171–188). Oxford: Oxford University Press.
- Feltz, A. (2007). The Knobe effect: A brief overview. *Journal of Mind and Behavior*, 3(4), 265–277.
- Kneer, & Bourgeois-Gironde. (2017). Mens rea ascription, expertise and outcome effects: Professional judges surveyed. *Cognition*, 169, 139-146
- Knobe, J. (2003). Intentional Action and Side Effects in Ordinary Language. *Analysis*, 63(3), 190- 194.
- Knobe, J., & Burra, A. (2006). The folk concept of intention and intentional action: A cross-cultural study. *Journal of Culture and Cognition*, 6, 113–132.
- Knobe, J., & Mendlow, G. (2004). The good, the bad, and the blameworthy: Understanding the role of evaluative reasoning in folk psychology. *Journal of Theoretical and Philosophical Psychology*, 24, 252–258.
- Nadelhoffer, T. (2004). On praise, side effects, and folk ascriptions of intentional action. *Journal of Theoretical and Philosophical Psychology*, 24, 196–213.
- Nadelhoffer, T. (2005). Skill, luck, control, and intentional action. *Philosophical Psychology*, 18, 341–352.
- Nichols, S., & Ulatowski, J. (2007). Intuitions and individual differences: The Knobe effect revisited. *Mind & Language*, 22(4), 346–365.
- Prochownik, K., Krebs, M., Wiegmann, A., & Horvath, J. (2020). CogSci 2020 Paper “Not as Bad as Painted? Legal Expertise, Intentionality Ascription, and Outcome Effects Revisited.” Retrieved from osf.io/n9h2b

An analysis of licensing of PPIs in negative imperatives in terms of implicit exemption from obviation

Makoto Kaneko (Aoyama Gakuin University, Japan)

1. Introduction

This study aims at investigating why, in negative imperatives, as in (1), whose predicate (e.g. *question somebody's immigration status*) denotes a non-intentional, PPI indefinites may be out-scoped by a clause-mate negation. The non-intentional nature of the predicate in (1) is confirmed by the compatibility with an adverb *inadvertently*.

- (1) [How to avoid offending minority customers?] Do not touch them. [...] Do not use profanity or racial slurs, and do not (*inadvertently*) question **somebody's** immigration status. (Coca)

After having examined Goncharov's (2018; 2020a,b) analyses of non-intentionality effect on negative imperatives and on licensing of PPIs, I advance an analysis which combines i) Szabolcsi's (2010) analysis in terms of implicit exemption from obviation and ii) Richardson's (1985) analysis of negative imperatives with non-intentional predicates.

2. Goncharov (2018; 2020a,b)

Goncharov (2018: 1) observes that in Slavic languages, negative imperatives are incompatible with perfective aspect, as in (2a), while this aspectual restriction "is lifted when the action expressed by the verb is interpreted as non-intentional", as in (2b). To account for this contrast, she assumes that intentional predicates are equipped with an intentional operator (**INT**), as in (3a), and that the combination of the assertion in (3b), the scalar implicature in (3d), and the presupposition due to the Intentional operator in (3e), amounts to a contradiction, which isn't, according to her, the case with non-intentional predicates in (2b). The presupposition in (3e) however doesn't seem to distinguish the intentional *open the window* and the non-intentional *fall down*: if 'falling down' starts, it also reached the end point in ordinary situations.

- (2)a. Ne otkryvaj / *otlroj okno! [Russian]
not open-IMPERF[ECTIVE].IMP[ERATIVE]/ open-PERF[ECTIVE].IMP window
'Don't open the window!'

- b. Ostorozhno! Ne upadi! [Russian]
careful not fall- PERF.IMP 'Be careful! Don't fall down!' (Goncharov 2018: 2)

- (3)a. Not **INT** open-PERF the window! [paraphrase of (2a)]
b. Assertion: You must not end up opening the window.
c. Alternative: You must not start to open the window.
d. Scalar implicature: \neg (You must not start to open the window) = You may start to open the window.
e. Presupposition due to **INT**: It must be the case that if you start to open the window, you end up opening it.

Goncharov (2020a,b) revisits Szabolcsi's (2004) observation that PPI indefinites in the controlled complement of *want* may be out-scoped by the matrix negation when the complement predicate is non-intentional, as in (4b), unlike in cases with intentional predicates, as in (4a). To account for this contrast, Goncharov (2020a) argues that the complement denoting a non-intentional action induces a propositional operator **LUCK**, as in (5a), which induces the presupposition, as in (5b), i) that the referent of the matrix subject thinks that in some possible world near by the actual world, an event contrary to what she wants may happen (e.g. 'offending someone', for (4a)), and ii) that PPIs are licensed in this upward-entailing context. This analysis reminds Larrivée's (2012: 894) analysis of (6a) where the proposition including *some* is "presupposed or activated" and "treated as a whole, to which the negative applies to reject it entirely."

- (4)a. I don't want to call **someone**. [$*not > someone$] (Szabolcsi 2004: 417, footnote 10)

- b. I don't want to offend **someone**. [$\sqrt{not} > someone$] (ibid.)

- (5)a. I want [**LUCK** not [_{TP} PRO to offend **someone**]] (paraphrase of (4b))

- b. Presupposition due to **LUCK**: $\exists w' \in \text{Near-by}(\text{Do}_{x_{\text{speaker}}, w})$ [offend (someone) (speaker)(w')]

- (6)a. John didn't say **something**. Jane did. [=It is not John that said **something**, but it is Jane.] (Larrivée 2012: 883)

- b. " $\lambda w. \text{John won the lottery}(w)$ " presupposes " $\exists w' \in \text{Near-by}(\text{Do}_{x_j, w})$ [\neg won-the-lottery(j)(w')]" (Goncharov 2020a: 16)

But since **LUCK** is induced in matrix clauses, as in (6b), PPIs should be out-scoped by a clause-mate negation, for example, in "John didn't offend someone", which is not the case. Goncharov (2018) and Goncharov (2020a,b) further advance essentially different accounts for non-intentionality effect on negative imperatives and on licensing of PPIs. It is not clear how her analyses apply to non-intentionality effect on PPI-licensing in negative imperatives, as in (1).

3. Szabolcsi (2010) and Richardson (1985)

Szabolcsi (2010) herself propose a different analysis of PPI licensing in (4b). This author first invokes Farkas's (1988) idea of *responsibility*, i.e. a relation holding between an individual *i* and a situation *s* just in case *i* brings *s* about

intentionally. She then observes that in Hungarian, when the complement of volition verbs expresses a non-responsibility situation, obviation may be explicitly exempt and a subjunctive complement becomes acceptable, as in (7a), where a PPI, *valakit* ‘someone’, in the complement is appropriately licensed by the clause-external negation.

- (7)a. Nem akarom, hogy lelőjek **valakit**. (Hungarian)
 NEGwant.1SG that shoot.SBJV.1SG someone.ACC
 ‘I don’t want that I (inadvertently) shoot **someone**.’ (Szabolcsi 2010: 7)
 b. I want for it not to be the case that I offend **someone**. (*idem*. 6)

Szabolcsi (2010) next suggests that, parallel to Hungarian (7a), the English infinitive in (4a) is reanalyzed, because of its non-responsibility nature, as a subjunctive clause, as in (7b), where *someone* is licensed by the clause-external negation.

I next refer to Richardson’s (1985) analysis of negative imperatives. This author claims that, in (8a), “not falling of the ladder is for the carpenter not an act in any useful sense of the word at all, but rather a state of affairs which he must try to maintain if he is going to accomplish whatever his positive intentions might be”. In other words, negative imperatives like (8a) do not convey prohibitions but “commands to monitor against the events denoted by the VPs” (*idem*.247). He proposes to capture this meaning by means of coercing of a monitoring predicate, *take care*, as in (8b).

- (8)a. Don’t fall off the ladder! (Richardson 1985: 247)
 b. TAKE CARE not to fall off the ladder. (*ibid.*) [paraphrase of (8a)]

4. Proposal

I now propose, partly modifying Szabolcsi’s and Richardson’s ideas, that the negative imperatives in (1) and (2b) are semantically reanalyzed as (9a) and (9b). According to this analysis, Goncharov’s presuppositional content due to LUCK is put into at-issue content. PPIs and perfective aspect are appropriately licensed since negation is clause-external.

- (9)a. TAKE CARE for it **not** to be the case [that you question **somebody**’s immigration status].
 b. TAKE CARE for it **not** to be the case [that you fall- PERF down].
 (10) $[\lambda x. x \text{ TAKE CARE } [\neg \exists p [p = \lambda w' (\exists y. \text{you question } y\text{'s immigration status in } w')]]] \in \text{To-Do-List}(\text{you})$

Furthermore, adopting an analysis of imperative as property-denoting, I semantically represent (9a) by (10), which says that taking care for the relevant situation not to happen is a property for the hearer to put into her To-Do-List. According to this hypothesis, coercion of monitoring predicate TAKE CARE is motivated by avoiding type-mismatch between property-denotation of the imperative and proposition-denotation of the subjunctive clause (cf. Grano’s (2015) analysis of non-controlled infinitive of *intend*, as in “John intends for Bill to leave.” in terms of coercion of the cause predicate BRING ABOUT). The externalization of negation may be motivated by a need to avoid a semantic incongruity of wide scope of non-intentionality over negation, as shown by the contrast between (11a) (where the adverb instantiating non-intentionality, *inadvertently*, takes wide scope) and (11b) (where it takes narrow scope).

- (11)a. ??Inadvertently, I offend **no one**. b. There is **no one** I advertently offend.

Two evidences come for the analysis in terms of implicit exemption from obviation : i) It allows to account for licensing of a PPI in the French purpose clause in (12b), in a parallel fashion to the explicit case of exemption in (12a); ii) Japanese auxiliary –TE SHIMAU is ambiguous between a completion reading (with intentional predicates) and an *advertently*-reading (with non-intentional ones), and an imperative with <non-intentional predicate+ TESHIMAU> is most naturally interpreted as a wish accompanying a subjunctive clause or as parallel to a third person imperative, as in (12c).

- (12)a.[...] je maîtrise suffisamment mes sorts de glaces *pour ne pas que je blesse **quelqu’un** par inadvertance !* (google)
 ‘I sufficiently control my ice trails in order not that I inadvertently hurt **someone**.’
 b. [...] *pour ne pas déranger **quelqu’un*** je suis parti m’asseoir sur les derniers bancs [...] (google)
 ‘In order not to disturb **someone**, I left to sit down on the benches in the last row’.
 c. korona-ni kakat-te simae. (google)
 covid-19-DATcatch-TE SHIMAU-IMPERATIVE
 ‘Catch inadvertently the covid-19!’ = (I wish that) You inadvertently catch the covid-19!

References: Farkas, D. 1988. On obligatory control. *Linguistics & Philosophy* 11; Goncharov, J. 2018. Intentionality Effect in Imperatives. *Proceedings of FASL 27*; Goncharov, J. 2020a. Intentionality in the grammar (http://juliegongcharov.com/docs/three_puzzles_draft.pdf); Goncharov, J. 2020b. Language and intentions. *GLOW* 43. <https://osf.io/zm8k3/>; Grano, T. 2015. Getting your to-do list under control. *NELS* 45; Larivée, P. 2012. Positive polarity items, negation, activated propositions, *Linguistics* 50; Richardson, J. 1985. Agenthood and Ease. *Proceedings from the 21st CLS*; Szabolcsi, A. 2004. Positive polarity negative polarity. *Natural Language & Linguistic Theory* 22; Szabolcsi, A. 2010. Infinitives vs. subjunctives: What do we learn from obviation and from exemptions from obviation? (<https://philpapers.org/rec/SZAIVS>)

Multiplicity and Culmination

Stefano Papa (independent researcher)

Key Concepts: *Routine, Action Logic, Recall, Retrieve*

1. Routines and distinctness

In English, the word routine is a french borrowing dating from the 1670s, a time when french terms in the military sphere were a sign of conceptual innovation rather than a victorious nation's imposing its language on neighbouring countries. (See Francis Bacon, *The Essays*, *On Studies*; an early description of procedural routines). The meaning of the word, customary way, beaten path, points in the direction of commerce and trade. Much later the computing meaning of the word shows up: A sequence of instructions for performing a task that forms a program or a distinct part thereof. One of the most interesting efforts at making the latter concept distinct is certainly the short contribution *Checking a Large Routine* (1949) by Alan Turing. These uses of the concept 'routine' exhibit some ambiguity: on the one side, a routine can be an intensional implicate of a habit, in the sense that having a worked-out part can help making a habit of something (for instance, the habit of performing the routine regularly); on the other side, a routine can be self-strengthening without resulting in a habit (at least not in the sense of outputting, say, religious or moral attitudes as a result). As a consequence, a habit can be both conceived of as a continuum of paths, a complete and thorough policy; and a collection of random walks, given a context where reliability and maximality fail (think of praying as an example for both: an effective way of training mindfulness given a pre-institutional setting; or an empty form of bigotry in a secularised context). In the context of a theory of (information in) action, an investigation of definitional markers of 'routine' is thus a promising one. In 1985a (*Synthese* 65, pp.185-210), Krister Segerberg introduces the concept of a routine as a primitive to be used as a basis for developing a philosophy of action: "To do something is to run a routine. To be able to do something is to have a routine available. To deliberate is to search for a routine" (Segerberg 1985a, p. 188). Markers of a routine are the concept of level (for instance, a chain of command, Segerberg's example), and that of perspective, being a partially ordered set of levels. The key function of 'routine' in Segerberg's account for action is that a routine provides, so to

speak, the action counting device for intentions. An agent intends to bring about a certain state of affairs S , thus, means that the agent selects an S individuating routine. In an assesment of Segerberg's approach to action theory, Dag Elgesem pointed out that some implicit assumptions of the concept of a routine result in inconsistencies (Elegesem Synthese 1990, vol. 85, No 1. pp 153-177). Elegesem shows that exactly the jointure routine-intention yields, in Segerberg's modeling, the unwanted result of a routine that never fails (Elegesem 1990, p. 163: Int $a \beta$ Real a).

2. Incompleteness and Causation

Segerberg's focusing on action incompleteness (Segerberg1992) can be seen as a reaction to Elgesem's criticism (as well as to other critical receptions by Thomason and Surendonk). The notion of action incompleteness, therefore, can be taken to be a refinement of 'routine'. The informal roots of 'routine' reach in the idea that acting means “to interfere with the course of nature” (v. Wright1963). An advantage of von Wright’s predicament ‘actions are not events’ is of considering events as structured (as encoded in linguistic representations) into an inner event (associated with the aspects of change and aspects) and an outer noneventive part (agentiveness and causation). The formal structure of acting, thus, is mirrored in descriptive categories as “to bring something about, to sustain something, to destroy something or suppress it” (v. Wright, Norm and Action.1963).

3. Sample, Multiplicity, Type Sensitivity

The idea this contribution intends to take stock with is that habit-forming parts (retention, learning, retrieval, indexing, refinement points for information) have to be incorporated in a model of acting in order for it to account for the eventive - non eventive divide within an acting sequence (act types and action tokens). Accordingly, going on ϕ -ing (performing a routine) is modeled as a multiplicity function over action tokens: keep ϕ -ing until token $\Phi_{i+1} \beta \Psi$, that is until Φ_i defines a multiset with Φ -multiplicity i over tokens Φ_n, Ψ_m . The intuitive reading of the above being that the informal concept of routine has, as one of its implicit aspects, the functional idea of interaction, that is of turn-taking. The incompleteness of actions is, in this respect, a trivial fact about routines with other than trivial consequences.

Letting structure speak with authority: Constraining agents' choices with French *laisser*

Clémentine Raffy*, Marta Donazzan°, Bridget Copley@, Klaus von Heusinger*

*Universität zu Köln °Université de Nantes/LLING @Université Paris8/SFL

Agent's actions and intentions can be prompted or hindered in multiple ways. Across languages, verbs that lexicalize the causative primitives of CAUSE, ENABLE or PREVENT (Wolff & Song 2003) can help us understand the nature of agency, precisely because they involve multiple participants which are sometimes seen as being in a position of influencing each other via different types of force relations. In this talk, we focus on the role of **authority**, intended as an influence that affects the choices available to a free agent with respect to the actions in service of their goal. We show that, while many causative verbs seem to imply the type of force relation between the participants in their lexical meaning, the French causative verb *laisser* is underspecified: the type of influence exerted by the two participants in a *laisser* relation is determined by the syntactic structure of the causative construction.

(i) Authority and choice restriction. What does it mean to have authority over an agent? To be an authority means to be able to influence the action of a free agent. In the semantic representation of syntactic causative constructions such as (1) and (2), we model authority as the ability to constrain the set of alternative actions available to the lower agent denoted by the non-finite predicate.

- (1) John prevented Mark from watching TV. (2) Lucy allowed Peter to go out tonight.
ALT_{MARK} {watch TV, not watch TV} ALT_{PETER} {go out, not go out}
John sets value of ALT_{MARK} := {not watch TV} Lucy sets value of ALT_{PETER} := {go out, not go out}

The subject of the lower clause in (1) has an alternative choice, which is determined by the denotation of the predicate (Mark can either watch TV *or* not watch TV); this alternative is restricted by the intention of the matrix subject John, and as a result the only available action is “not watch TV”. John's intention is therefore restricting Mark's freedom of action. Conversely, in ENABLE situations, the first participant's intention does not affect the second participant's choice between alternative actions. In (2), the set of alternative actions ALT available to Peter includes the options of going out and not going out, and Lucy's intentions do not restrict the set. Crucially, however, in both the configurations expressed by *prevent* and *allow*, the agentive subject of the matrix causative verb is understood as having the ability to restrict the set of choices available to the other participant(s): if Lucy is in a position to *allow*, she has enough authority to *prevent* too.

(ii) Structural constraints on authority relations. The French causative verb *laisser* (‘let’) is generally described as realizing an ENABLE relation. The syntactic causative can be realized in either of two ways: the embedded subject can appear before the non-finite predicate (a pre-V construction (3a)) or after it (a post-V construction (3b)).

- (3) a. Jean a laissé Pierre manger. b. Jean a laissé manger Pierre.
John let.PF Pierre eat.INF Jean let.PF eat.INF Pierre
‘Jean let Pierre eat.’

Kayne (1975:222) points out that the two structures correlate with two interpretations: (4a) can be interpreted as a situation where the guard acted with “complicity” or “deliberate neglect” with respect to the prisoner's attempt to escape, while there is no such inference in (4b). In other words, we could say that (4a) carries an *intentional flavor*: it seems to be the guard's intention to give to the prisoner the choice to escape. No such inference arises in the post-V construction (4b).

- (4) a. Le gardien a laissé le prisonnier s'échapper.
b. Le gardien a laissé s'échapper le prisonnier. (Kayne 1975:222)
‘The guard let the prisoner escape.’

But where does this *intentional flavor* come from? We argue that i) (4a) necessarily makes reference to the authority relation between the two participants, and ii) an authority relation necessarily involves the intention of the authority. Conversely, (4b) does not make any reference to an authority relation and thus does not trigger the interpretation of an intentional agent.

To see this, consider the following contrast:

- (5) a. La juge a laissé l'accusé parler.
'The judge let the defendant speak.'
- b. La juge a laissé parler l'accusé.
'The judge let the defendant speak.'
- (6) a. #L'accusé a laissé la juge parler.
'The defendant let the judge speak.'
- b. L'accusé a laissé parler la juge.
'The defendant let the judge speak.'

In (5), we have kept the same (natural) authority relation as in (4): the subject of *laisser* has authority over the *lower agent*: both sentences are felicitous. However, in (6), we reversed the authority levels: it is infelicitous for the defendant to have authority over the judge, and this accounts for the infelicity of the pre-V construction. We submit that in the case of *laisser* the role of authority is determined by the possibility of constructing a syntactic structure where the lower subject is realized as the agent of the embedded verb. In other words, in pre-V constructions, *laisser* embeds a clause whose subject is interpreted as an Agent (possibly by a VoiceP, Alexiadou 2014). In this first case, the Agent introduces the ALT set, which the matrix subject must be able to restrict. Post-V constructions, on the other hand, feature a vP complement: the interpretation is that of a no-choice: an action that has already started.

- (5') La juge a laissé l'accusé parler. (5'') L'accusé a laissé parler la juge.
- a. LAISSER [_{VoiceP}] a. LAISSER [_{vP}]
- b. Judge [ALT_{defendant} {speak, NOT speak}] b. Defendant [judge speak]

Assuming that alternatives are propositional and therefore can take negation, whereas event descriptions cannot, the structure can account for the contrast in the use of negation in the complement of *laisser* – indeed, lower negation can be inserted in pre-V (7a) but not in post-V (7b):

- (7) a. Marc laisse ses enfants ne pas travailler. b. *Marc laisse ne pas travailler ses enfants.
Marc let.PRS his children NEG work.INF. Marc let.PRS NEG work.INF his children
'Marc let his children not work.'

Additionally, the presence of available alternatives can be signaled through Free Choice Items (FCIs) like *n'importe-quoi* 'any' (Staraki 2018). In (8a), the passenger has authority over the driver (of a taxi) and chooses not to restrict the choice of the driver to take any exit he pleases. (8b) on the other hand sounds odd: there is a clash between the absence of alternatives in the structure (post-V) and the presence of an FCI.

- (8) a. Le passager a laissé le conducteur prendre n'importe quelle sortie.
The passenger let.PF the driver take.INF FCI exit
- b. ?Le passager a laissé prendre n'importe quelle sortie au conducteur.
The passenger let.PF take.INF FCI exit PREP the driver
'The passenger let the driver take any exit.'

N'importe-lequel forces alternatives, which forces a lower free agent, which forces authority, which forces a pre-V structure, and disallows the post-V structure.

To conclude, in this paper we probed the relation between authority and intentional causation. We defined authority as the influence that an agent has over the action of another free agent, modeled as the restriction on a set of alternative propositions. We have shown that authority relation can be represented structurally – when there are two Agents involved, causative verbs like *laisser* can yield two interpretations based on the structure they embed.

REFERENCES

- Alexiadou, A. (2014). Active, middle, and passive: the morpho-syntax of Voice. *Catalan Journal of Linguistics*, 13, 19-40. • Staraki, E. (2018). Free choice inference in imperatives: The preference allocation hypothesis. *Lingua*. • Kayne, R. S. (1975). *French syntax: The transformational cycle*. Cambridge, MA: MIT Press. • Vlachou, E. (2007) *Free choice in and out of context: semantics and distribution of French, Greek and*

English free choice items. LOT • Wolff, P., & Song, G. (2003). Models of causation and the semantics of causal verbs. *Cognitive Psychology*, 47(3), 276-332.

THE ROLE OF AGENCY IN CAUSAL SELECTION

Elitzur Bar-Asher Siegal¹, Noa Bassel¹ and York Hagmayer²

1. Hebrew University of Jerusalem; 2. Universität Göttingen

SETTING THE SCENE: Since the philosopher David Hume, causality is generally assumed to be a binary relation between a cause and an effect. It is evident, however, that the occurrence of any particular effect depends on the realization of a **set of conditions**. For example, the effect of a house burning down may depend upon a discarded cigarette bud, but also on oxygen, flammable material and the absence of fire fighters. Consequently, expressing causal statements based on a single condition involves **Causal Selection**, i.e. teasing apart causes from mere background/enabling conditions. Accounts of causal selection maintain that causes and conditions hold similar logical relationships to the effect (necessity/counterfactuality (Lewis 1973)) and therefore, the choice of *the* cause should be accounted for via other types of criteria, such as normality (Icard, Kominsky & Knobe 2017 *inter alia*), or knowledge/interest based conversational principles (Beebe 2004 *inter alia*).

Following insights from Dowty (1979), we argue that causal selection is a linguistic phenomenon, under which causative expressions pose restrictions on what can appear as “the cause”. In other words, we rephrase the philosophical question of causal selection as a linguistic puzzle, asking **for each condition under what terms can it be represented as the cause in a given causative construction**. This question relates to the ongoing effort by linguists to capture the semantics of lexical causative verbs (*open*, *break*) vs. periphrastic causative constructions (e.g. *cause to*), (Fodor 1970, Neeleman & Van de Koot 2012, Maienborn & Hertfelder 2017, Lauer & Nadathur 2020, *inter alia*). It has been noted since at least Hall (1965) that lexical and periphrastic causatives have different semantics, given that configurations that are close paraphrases of one another such as (1-2) exhibit asymmetrical entailment relations:

(1) *Mary closed the door.* \models *Mary caused the closing of the door.*

(2) *Mary caused the closing of the door.* $\not\models$ *Mary closed the door.*

This contrast is commonly ascribed to the lexical causative having a prerequisite of “direct causation” (Fodor 1970, Katz 1970, Shibatani 1976, Wolff 2003, Martin 2018, Bar-Asher Siegal & Baglini 2020 *inter alia*). It is in this context that the role of agency becomes significant, since among the various characteristics of direct causation, the notion of agency was invoked. Cruse (1979), for example, argues that “we must understand ‘direct’ to mean that no agent intervenes in the chain of causation between the causer (represented by the subject of the verb) and the sufferer of the effect (represented by the object)”. This claim should be contextualized in the association, often made among linguists, between causation and agency, as for example DeLancey (1984) states that “ultimate cause can only be an act of volition on the part of a (thus defined) prototypical agent” – and it has been repeatedly claim that causation/agency plays a significant role in the grammatical relations represented by the syntax.

GOALS: We use an experimental approach to examine the role of linguistic construction in causal selection, and its interaction with temporal order and various properties considered in the psychological and philosophical literature as factors for normative conventions – and therefore as contributors to causal selection. In this context we examine to what extent agency plays a significant role in determining the identity of “the cause” in causative expressions.

DESIGN: In a series of four experiments, participants were presented with scenarios in which two causes conjunctively generate an effect, i.e., both causes were individually necessary and jointly sufficient for the effect to occur. Participants were asked to rate, on a scale of 1-7, the level of adequacy of two types of causal statements: one featuring a lexical causative (e.g., *Mary opened the door*) and the other a periphrastic causative (e.g., *Mary caused the door to open*). The linguistic construction was the primary factor, with each trial manipulating an additional factor: (i) Temporal order (Exp.1-4); (ii) Event vs. Agent (Exp 2) ;(iii) deviation from social norms (Exp.3); (iv) Foreseeability (Exp.4). All experiments were designed in English and conducted online. Samples sizes for each experiment were chosen to yield an adequate power for a moderate effect of causative construction.

RESULTS: Figure 1 shows mean ratings of the causal statements and their confidence intervals. Temporal order had an effect on both types of constructions, contra the common claim relating direct causation only with lexical causatives. Norm violation and foreseeability showed variable interaction with construction and order. Event/Agent manipulations did not affect judgments when the Agent could not have anticipated the effect, suggesting that when it comes to causal selection, agency can be reduced to foreseeability. Tables 1 summarizes the interaction of order, norm violation and foreseeability with linguistic construction.

DISCUSSION: The results show that speakers’ evaluations of the adequacy of different causal statements vis à vis a particular state-of-affairs vary systematically, depending on the type of linguistic expression employed to describe them. This variation indicates that causal selection depends on linguistic facts (i.e. the choice of constructions) and not merely on the metaphysical or cognitive characteristics of the relata. While these

findings are in line with the “direct causation” analysis of lexical causatives, the effect of temporal order on periphrastic causatives is unexpected. Following the reported results, as well as further ongoing trials, we suggest to revise the constraints on both types of constructions with respect to causal selection, as follows: For *cause to* constructions, the higher sensitivity to norm-violation (trial 3) and the ability of the participants to foresee the effect (trial 4) pertains to the degree of **responsibility** attributed to the condition *wrt* to the effect. In addition, an event is perceived as responsible for the effect, if it is the last to complete the set of sufficient conditions. With respect to the lexical construction, we define two modes of completion of the sufficient set: **Objective take**, where the last event to complete a sufficient set is the cause of the effect. **Subjective take**, in which the last event that is unexpected by involved participants is perceived as the cause (foreseeability).

So far these results demonstrate that the selection of events and agents as “the cause” of the causative constructions is based on similar principles. An ongoing trial examines casual selection when only one condition in a set is an agent, and the extent this factor affects judgments comparing to the parameters tested in trials 1-4.

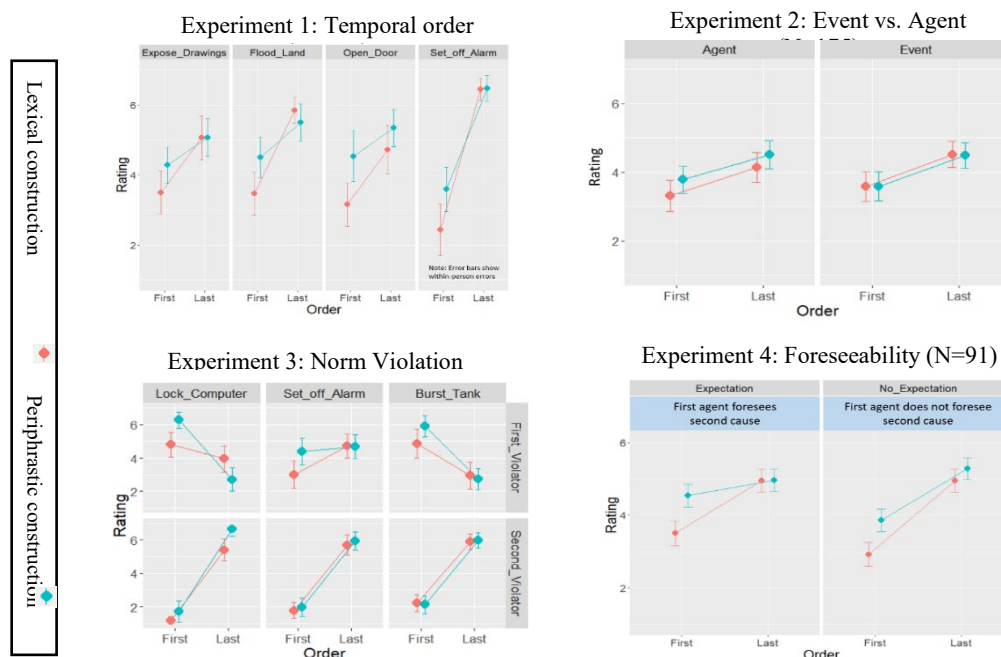


Figure I: Compatibility judgment per cause and linguistic construction in Trials 1-4

	Lexical construction	=	Periphrastic construction (<i>cause to</i>)
Order (completion of a sufficient set)	Always a factor	=	Always a factor
Violation of Norms	Sometimes a factor	>	Always a factor
Foreseeability	Always a factor	>	Always a factor

Table 1: Interaction of factors and linguistic constructions

BIBLIOGRAPHY: Baglini & Bar-Asher Siegal. 2020. Direct causation: A new approach to an old question; Beebee. 2004. Causing and Nothingness; Cruse. 1972. A note on English causatives. Linguistic Inquiry; DeLancey S. 1984. Notes on agentivity and causation; Fodor 1970. Three Reasons for Not Deriving "Kill" from "Cause to Die"; Hall. 1965 *Subject and Object in English*; Icard, Kominsky & Knobe. 2017. Normality and actual causal strength; Katz. 1970. Interpretive Semantics vs. Generative Semantics; Lauer & Nadathur. 2020 Causal necessity, causal sufficiency, and the implications of causative verbs; Lewis. 1973. Causation. *JoP* 70; Maienborn & Herdtfelder. 2017. Eventive vs. stative causation: The case of German causal von-modifiers. Linguistics and Philosophy; Martin. 2018. Time in probabilistic causation: direct vs. indirect uses of lexical causative verbs. Mill. 1884. *A system of logic*; Neeleman & Van de Koot. 2012. The Linguistic Expression of Causation; Pearl. 2000. *Causality: Models, reasoning, and inference*; Wolff. 2003. Direct causation in the linguistic coding and individuation of causal events; Shibatani 1976. The grammar of causative constructions: A conspectus.

Renouncing the attempt versus perpetration distinction

Izabela Skoczeń (Jagiellonian University) and Markus Kneer (University of Zurich)

John lights a long fuse in order to burn his enemy's house. After a short while, John has second thoughts. Consequently, he tries to stamp out the fuse. There are two possible outcomes. The first outcome is lucky: John stamps out the fuse easily and the enemy survives. However, it could also be that a strong wind prevents John from stamping out the fuse. As a result, John's enemy dies anyway.

In both cases, the lucky and the unlucky one, John has the same quality of the will: he initially intends a criminal outcome, yet later, on second thoughts, desists from his criminal intent and does everything he can to prevent the criminal outcome. While his actions are intentional, the outcome is accidental as it is not fully under John's control – it is subject to causal luck.

If one believes that it is one's quality of the will rather than an accidental outcome of an action which should determine, for instance moral and legal, responsibility, then irrespective of the outcome of John's action, he should be held equally responsible in both cases. Moreover, if quality of the will is to be decisive for responsibility, then John is surely more culpable than someone who never even undertook lighting a fuse in order to kill an enemy. However, John is less culpable than someone who would not entertain second thoughts and would not even try to stamp out the lit fuse. This stance could be labeled internalism about responsibility and culpability, as it goes along Kantian lines of reasoning.

By contrast, if one thinks that external factors, such as for instance the outcome of one's action, should determine one's responsibility or culpability, then John in the lucky case is less culpable or responsible than John in the unlucky case, as only in the latter the enemy dies. However, given that the outcome of John's action is not under his control, but is subject to outcome and causal luck, an externalist, consequentialist approach to culpability and responsibility is unfair.

If one looks at the general rules governing criminal intention (*mens rea*) ascriptions in legal systems around the globe, these rules are largely Kantian (internalist), rather than consequentialist (or externalists). After all intention is a necessary prerequisite of culpability. However,

surprisingly, both common and civil law systems treat the lucky and unlucky cases differently. If John is lucky and stamps out the fuse, John commits merely an attempt (one can 'attempt' only if one fails to achieve the goal). Consequently, he can use the so-called renunciation defense in court. This defense consists of a mitigation of punishment due to the fact that John completely and voluntarily desisted from his criminal enterprise. By contrast, if John is unlucky and the fuse cannot be stamped out due to the harsh wind, then John is taken to be a perpetrator, he cannot resort to the renunciation defense, even though he did everything he could to stamp out the fuse and, as a result, receives the full punishment. Yet are these rules the expression of reflective, all things considered views on legal responsibility? Or rather, are they the result of a bias arising out of the fact that in real life we never see the relevant counterfactual? In other words, juries and judges see either the lucky or the unlucky John separately, they never compare the two cases.

In order to test the bias hypothesis, we performed a series of experiments (total N=800). We employed a between subject and a within subjects' experimental design. In the between subjects' design, participants are presented with either a lucky or an unlucky renunciation case. Next, participants are asked questions about moral and legal responsibility (blame, wrongness of action and punishment), probability of the outcome occurring (subjective and objective) and mental states. We also performed the same experiment in a within subjects' design.

We found that (i) outcome influences moral and legal responsibility ascriptions; (ii) the influence of outcome is mitigated in the within subjects design as compared to the between subjects design (effect size of the difference in between subjects $d = .87$; versus effect in within subjects $d = .35$); (iii) probability mediates the relation between outcome and mental state ascriptions.

Based on our results, we point toward the conclusion that the folk concept of fair and just responsibility requires that agents are held responsible only for outcomes that are under their control. Since legal rules on the renunciation defense in common and civil law systems run counter to this claim, they should be reformulated.

References:

Alexander, L., Ferzan, K. K., & Morse, S. J. (2009). *Crime and Culpability: A Theory of Criminal Law*. Cambridge University Press.

- Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, 126(4), 556–574
- Kneer, M., & Machery, E. (2019). No luck for moral luck. *Cognition*, 182, 331–348.
- Knobe, J. (2004). Intention, intentional action and moral considerations. *Analysis*, 64(2), 181–187.
- Knobe, Joshua, & Shapiro, S. (forthcoming). What Cognitive Science Can Teach Us about Proximate Causation.

Agentivity (and lack thereof) in Ibero-Romance middles: evidence for *i**

Imanol Suárez-Palma

University of Florida

Middle constructions in Ibero-Romance are characterized by the presence of the reflexive clitic *se*, imperfective tenses, agreement between the verb and the preverbal definite DP theme, and the impossibility of introducing an agent by means of a *by*-phrase. Moreover, these sentences denote the participation of a generic external argument in the event crosslinguistically (Ackema & Schoorlemmer 2006, *inter alia*). Interestingly, descriptive work on Asturian (ALLA 2001) notes that the reflexive appears to be optional in middles (1a), unlike in Spanish (1b) or Catalan (1c).

- (1) a. Les fueyes mueyen/muéyense col orbayu (*por Xuan). (Asturian)
the leaves wet wet.REFL with.the light rain by Xuan
b. Las hojas *(se) mojan con la lluvia ligera (*por Juan). (Spanish)
the leaves REFL wet with the rain light by Juan
c. Les fulles *(es) mullen amb la pluja lleugera (*per Joan). (Catalan)
the leaves REFL wet with the rain light by Joan
'Leaves get wet/are wetted with the light rain.'

However, such alleged optionality in Asturian is only possible with verbs denoting a change of state/location entering the causative alternation; predicates denoting activities or accomplishments necessarily require the reflexive (2).

- (2) a. Les noveles de misteriu lléen*(se) con facilidad. (Asturian)
the novels of mystery read.REFL with ease
b. Las novelas de misterio *(se) leen con facilidad. (Spanish)
the novels of mystery REFL read with ease
'Mystery novels read easily.'

Additionally, the reflexive clitic, although optional, is occasionally required with certain unaccusative verbs in languages like Spanish to denote telicity (Cuervo 2014); in Asturian, nonetheless, it is banned.

- (3) a. Los guajes cayen(*se) fácil cuando entamen a andar. (Asturian)
the kids fall.REFL easy when begin.1PL to walk
b. Los niños (se) caen fácilmente cuando comienzan a caminar. (Spanish)
the kids REFL fall easily when begin.1PL to walk
'Kids fall down easily when they start walking.'

Here, I propose that the presence of the reflexive in middle constructions in these languages is associated with the projection of a passive Voice head encoding the participation of a generic external argument in the event (Kratzer 1996; Schäfer 2008). In fact, this argument can control into purpose clauses, which is taken as evidence for the presence of an implicit agent (Bhatt & Pancheva 2006), so long as the reflexive clitic is present, as shown in (4).

- (4) a. Esti material rúempe*(se) fácil *pa* reciclalo *afechiscamente*. (Asturian)
this material breaks.REFL easy for recycle-it.ACC adequately
b. Este material *(se) rompe fácilmente *para* reciclarlo *adecuadamente*. (Spanish)
this material REFL breaks easily for recycle.it.ACC adequately
Intended: 'This material is easy to break in order to recycle it adequately.'

These sentences also allow a non-agentive reading, which can be enhanced by means of the insertion of a PP like *por sí mismo/mesmu* ('by itself'). Crucially, such PP in Asturian can only be licensed in the absence of the reflexive clitic.

- (5) a. Esti material ruempe*(se) fácil *por sí mesmo*. (Asturian)
this material breaks.REFL easy by itself
b. Este material *(se) rompe fácilmente *por sí mismo*. (Spanish)

this material REFL breaks easily by itself
 ‘This material breaks easily by itself.’

Thus, I claim that two possible configurations can yield middle sentences containing change-of-state predicates: (i) a generic *se*-passive (4) containing a passive Voice head and a bieventive structure comprising an activity subevent and a stative one ($\nu_{DO} + \nu_{BE}$) (Cuervo 2003); and (ii) a generic inchoative, without Voice, and composed by a subevent of change and a stative one ($\nu_{GO} + \nu_{BE}$). While the reflexive clitic in (4) is the spell-out form of the passivized Voice head, in (5b) it spells out the subevent of change (ν_{GO}); the difference between Spanish –as well as other Ibero-Romance languages, such as Catalan– and Asturian, is the fact that ν_{GO} in the latter does not have a phonetic form.

Moreover, middle contexts containing change of state verbs allow the insertion of an affected applicative head (Cuervo 2003; Suárez-Palma 2020) introducing a non-core dative argument, which can be interpreted either as affected by the theme’s resulting state if merged below the first subevent (either ν_{DO} or ν_{GO}), or as unintentional causer of the event if merged on top. Crucially, the later interpretation is only available in the generic inchoative variant (6a), where the position above ν_{GO} is not filled by Voice, i.e. in the generic *se*-passive counterpart (6b).

(6) *Generic inchoative*

- a. A Xuan,_i esti material ruémpe-y_i fácil. (Asturian)
 to Xuan.DAT this material breaks-3SG.DAT easy
 i. ‘Xuan accidentally causes this material to break easily.’
 ii. ‘This material breaks easily, and Xuan is affected by it.’
 iii. ~~‘It is easy to break Xuan’s material, and he is affected by it.’~~

Generic se-passive

- b. A Xuan,_i esti material ruémpe-se-y_i fácil. (Asturian)
 to Xuan.DAT this material breaks.REFL-3SG.DAT easy
 i. ~~‘Xuan accidentally causes this material to break easily.’~~
 ii. ~~‘This material breaks easily, and Xuan is affected by it.’~~
 iii. ‘It is easy to break Xuan’s material, and he is affected by it.’

These data suggest there exists a competition between Appl and Voice for the position atop the first subevent. I propose that such phenomenon can be explained by adding to Wood and Marantz’ (2017) notion of *i**, i.e. an argument-introducing functional head whose spell-out form varies depending on its surrounding environment; thus, *i** would be realized as Voice in the context of ν_{DO} and as Appl when its complement is ν_{GO} . The derivations for (6) are shown in (7).

(7) *Applicative as unintentional causer (no Voice)*

- a. [_{AppIP} [_{DP} **a Xuan**]_i] [_{Appl} -y_i] [_{PGO} [_{GO} [_{PBE} [_{DP} *esti material*] [_{BE} [_√ *romp*-]]]]]]]
Applicative as affected by an event (with or without Voice)
 b. ([_{VoiceP} [_{Voice} **se**] [_{PDO/GO} [_{DO/GO} [_{AppIP} [_{DP} **a Xuan**]_i] [_{Appl} -y [_{PBE} [_{DP} *esti material*] [_{BE} [_√ *romp*-]]]]]]]]]]]

In sum, this proposal demonstrates that a generic passive and a generic inchoative can underlie middle constructions containing a change of state predicate in Ibero-Romance, giving rise to syntactic ambiguity in those languages where *se* spells out both Voice and ν_{GO} ; additionally, it shows that the presence of the reflexive in Asturian middle sentences is exclusive to those predicates obligatorily subcategorizing for an external argument; finally, the three different interpretations dative arguments can have in these contexts supports Wood & Marantz’s *i** hypothesis.

References:

ALLA. 2001. *Gramática de la Lengua Asturiana*. Oviedo: ALLA. Ackema, P. & M. Schoorlemmer. 2006. Middles. In M. Everaert, H. van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, Vol. 1, 131-203. Oxford: Blackwell. Bhatt, R. & R. Pancheva. 2006. Implicit Arguments. In M. Everaert, H. van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, vol. 1, 558–588. Oxford: Blackwell. Cuervo, M.C. 2003. *Datives at Large*. Cambridge: MIT dissertation. Cuervo, M.C. 2014. Alternating unaccusatives and the distribution of roots. *Lingua* 141. 48–70. Kratzer, A. 1996. Severing the external argument from its verb. In J. Rooryck & L. Zaring (eds.), *Phrase structure and the lexicon*, 109–137. Springer, Dordrecht.

Schäfer, F. 2008. *The syntax of (anti-)causatives: External arguments in change-of-state contexts.* Amsterdam: John Benjamins Publishing. **Suárez-Palma, I.** Applied arguments in Spanish inchoative middle constructions. *Glossa* 5(1), 9. **Wood, J. & A. Marantz. 2017.** The interpretation of external arguments. In R. D'Alessandro et al. (eds.), *The Verbal Domain*, 255–278. Oxford: OUP.

Has the Side-Effect Effect Been Cancelled? (No, Not Yet.)

Justin Sytsma, Victoria University of Wellington
John Schwenkler, Florida State University
Robert Bishop, Florida State University

In the *Side-Effect Effect* (SEE), side effects that are regarded as morally bad are judged as more intentional than side effects that are neutral or morally good (Knobe 2003a). Investigation of the SEE has replicated the effect with different cases (Knobe 2003b, Nadleoffer 2004, Knobe 2006, Cushman & Mele 2007), languages (Mizumoto 2018, Knobe & Burra 2006), and concepts (Beebe & Buckwalter 2010, Pettit & Knobe 2009).

There is, however, persistent disagreement about the explanation of the SEE. The first, most straightforward account is that the effect arises because participants mean just what they say: the effect reflects their *beliefs* about whether certain things are done intentionally or not. But an alternative account claims that the asymmetry arises due to pragmatic pressure to express moral censure. The pragmatic account holds that, since a person is blameworthy for the harmful side-effects of their actions in a way that they are not praiseworthy for side-effects that are beneficial, participants express agreement with the statement that the harmful side-effects were intentional in order to avoid implying that they are excusing them, whereas the same pressure is not felt in connection with neutral or beneficial side-effects. If the pragmatic account is correct, it should be possible to cancel the SEE by giving participants a different way of attributing moral responsibility.

A forthcoming paper by Lindauer and Southwood (hereafter ‘L&S’) purports to do just this. In the crucial condition of their study, participants read the harm version of Knobe’s (2003a) chairman vignette and then rated the following *cancelling* statement on a 7-point scale:

- (C) The chairman didn’t intentionally harm the environment, but he knowingly harmed the environment, and he is morally responsible and should be blamed for doing so.

In the other conditions, participants read either the help or harm version of the chairman vignette and rated one of the following *simple* statements instead:

- (S_{help}) The chairman didn’t intentionally help the environment.
(S_{harm}) The chairman didn’t intentionally harm the environment.

L&S’s main finding was that, while ratings of (S_{help}) and (S_{harm}) exhibited the usual SEE, their participants agreed with (C) to about the same extent as they agreed with (S_{help}). That is, participants who read the version of the vignette in which the chairman had *harmed* the environment agreed just as much with (C), which says that this harm was unintentional, as participants who read the version in which the chairman had *helped* the environment agreed with

(S_{help}). L&S take this finding to provide evidence for the pragmatic account of the SEE. Our paper shows why it does not.

To begin, notice that L&S's cancelling statement (C) has two parts, which are separated by the contrastive conjunction 'but'. The first part of (C) is the same as the statement (S_{harm}) above, which denies that the chairman harmed the environment intentionally. And the second part of (C) is a positive attribution of moral responsibility to the chairman:

(R) The chairman knowingly harmed the environment, and he is morally responsible and should be blamed for doing so.

According to L&S, their participants agreed with (C) because they agreed independently with both (S_{harm}) and (R), and the opportunity to express their agreement with the latter statement relieved what had been merely pragmatic pressure to deny the former. To explore whether this is the correct account of these findings, we conducted several experiments to test a further prediction that follows from L&S's account, namely that participants who are allowed to censure the chairman should prefer saying that he did *not* harm the environment intentionally over saying that he *did* harm the environment intentionally—since the first thing is supposed to be what they really believe. That is, Lindauer and Southwood's pragmatic account predicts that *participants who are given the opportunity to express strong moral censure of the chairman, by affirming a statement like (R), should tend to disagree overall with (S_{harm})*. In each of our experiments, this prediction was not borne out.

In our first study, participants read the harm version of the chairman vignette and then rank-ordered four variants of L&S's cancelling statement (C), including the following:

- (C1) The chairman intentionally harmed the environment, *and* he knowingly harmed the environment, is morally responsible for doing so, and should be blamed for it.
- (C2) The chairman did **not** intentionally harm the environment, *but* he knowingly harmed the environment, is morally responsible for doing so, and should be blamed for it.

As we have seen, L&S's pragmatic account predicts that participants should rank (C2) higher than (C1): for the final three clauses of each statement remove any pragmatic pressure to censure the chairman, and in this context participants should feel free to say that he did not harm the environment intentionally. But we found just the opposite, as 65.7% of our participants ranked (C1) higher than (C2). This pattern was statistically significant: $\chi^2=5.97, p=.015$.

In our second study, participants read the same vignette and then rated the four statements from our first study on a 7-point scale. Again, L&S's account predicts that (C2) will elicit higher ratings than (C1), at least when participants are given the opportunity to rate both of the statements. Again, we found the opposite, as mean ratings of (C1) were significantly higher than mean ratings of (C2): $t(195.76)=1.80, p=.037, d=.25$. This same pattern was observed for both a

between-participants design, in which each participant rated only one statement, and a within-participants design, in which each participant rated all four.

Finally, in our third study participants read the same vignette and then first indicated their level of agreement with the responsibility-attributing statement (R), after which they indicated their level of agreement with the simple statement (S_{harm}), which denies that the chairman harmed the environment intentionally. Again, L&S's account predicts that in this condition participants should tend to disagree with (S_{harm}), since the opportunity to express agreement with (R) removes pragmatic pressure to censure the chairman and allows them to say what they really think, namely that he did not harm the environment intentionally. And once again we found just the opposite, as mean ratings of (S_{harm}) were significantly above the neutral point: $t(53)=6.61$, $p<.001$, $d=.90$.

Taken together, these findings suffice to refute Lindauer and Southwood's account of their (forthcoming) findings. While participants in their experiment expressed higher agreement with their cancelling statement (C) than with their simple statement (S_{harm}) this was *not* simply because they agreed with what (S_{harm}) says on its own. The Side-Effect Effect remains uncanceled.

**Intentional Inaction, Blame, and Obligation:
A Case Study of “Ought Implies Can” Using Qualitative Methods**
By Kyle Thompson, PhD
Adjunct Instructor of Philosophy and Writing, Harvey Mudd College

In recent years, experimental philosophers have used quantitative surveys to explore a number of folk concepts related to moral judgment, such as intentional action, blame, and obligation. At the intersection of a number of these concepts is “ought implies can” (OIC), a principle often attributed to Kant which holds that moral obligation applies only to actions that an agent can carry out.¹ Unsurprisingly, experimental philosophers have taken their survey methods to OIC, including a prominent and rigorous study by Chituc, Henne, Sinnott-Armstrong, and De Brigard (Chituc et al.) in which an agent’s blameworthiness was varied to essentially stress test participants’ judgments of “ought” as it applied to an action rendered impossible.² Specifically, Chituc et al. argue that folk judgments of “ought,” which appeared swayed by an agent’s blameworthiness, don’t support that “ought” implies “can” based on the responses to two kinds of vignettes they administered: 1) scenarios in which agent was unintentionally unable to make a promised meeting (low blame), and 2) scenarios in which agent intentionally waited too long such that he was unable to make a promised meeting (high blame).³ A number of other experimental studies and analyses in this literature support the contention that OIC isn’t strongly upheld in folk cognition,^{4, 5, 6, 7, 8} while others challenge this contention in one way or another.^{9, 10, 11, 12, 13} This debate has sparked a fruitful discussion on folk concepts, OIC, and the challenges in capturing relevant folk judgments.

However, as with many studies in experimental philosophy (x-phi) which rely heavily or exclusively on quantitative methods, it is unclear whether the Chituc et al. study¹⁴—and the surrounding x-phi literature on OIC—provides clear evidence with which to evaluate participant usage of folk concepts. This results from an interpretive worry and a related methodological worry. The interpretive worry is whether participants’ selections on quantitative surveys reflect their judgments—i.e., whether a given survey is valid—and it is exacerbated by the methodological worry that quantitative surveys are a blunt and inaccurate measuring device for getting at participants’ judgments on concepts that are sought after precisely because they are messy and multifarious—i.e., OIC is debated in philosophy because it’s complicated. To address these worries as they pertain to OIC in particular, as well as x-phi literature in general, I modified one of Chituc et al.’s experiments by adding a layer of qualitative data via both the think aloud method¹⁵ and the qualitative interview.

Specifically, I had participants talk aloud as they completed surveys from Chituc et al. while I recorded their utterances; subsequently, I interviewed them about their experiences completing the surveys, again while recording. In this way, I could not only compare a given participant’s quantitative selections to his or her qualitative utterances—thereby running a kind of validity check on the survey—but I could also capture any interesting contours of folk judgments regarding ability, intentional action or inaction, blame, and obligation. In total, I found that participants’ judgments were significantly different than their quantitative selections indicated—interestingly, sometimes their utterances were the opposite of what their quantitative selections suggested. I also found that their judgments were overall more complicated, messy and multifarious than the quantitative surveys had implied. Conflicting with Chituc et al.’s results, participants in large part upheld or preserved OIC rather than violated it even when given ample opportunities to do so. Often times, while participants’ quantitative results suggested a violation of OIC, their utterances, now given open-ended space for clarification and explanation, said otherwise.

For this workshop, I plan to present some of the more salient and interesting findings within my qualitative data—after a brief overview of my quantitative results—in order to argue that my results have a number of important implications for understanding folk concepts and their relation to debates in analytic philosophy. First, my study suggests that quantitative surveys alone are a poor guide to folk

judgments in OIC unless they can be designed to account for complex features of folk judgments that are difficult to appreciate without first employing qualitative methods. That is, I'll suggest experimental philosophers embrace a triangulation of both quantitative and qualitative methods. Second, I'll argue that it's probable that studies in x-phi outside of OIC, given the high prevalence of survey methods in x-phi generally, suffer from similar methodological limitations that present skewed interpretations of folk judgments. Again, this suggests the value of triangulation of methods in x-phi. Lastly, my study suggests that folk judgments relating to intentional action or inaction, blame, and obligation are far more complicated and multifarious than they are often treated in the x-phi literature on OIC. In particular, ascriptions of "ought" are sensitive to concreteness (temporal concreteness in particular), the perceived control an agent has over present and past actions, the role of moral residue, and more. To conclude I'll discuss the limitations of my own study and suggest what future research might explore.

References

- Andow, James. "Qualitative Tools and Experimental Philosophy." *Philosophical Psychology* 29, no. 8 (2016): 1128-1141. Accessed December 22, 2016. doi: 10.1080/09515089.2016.1224826.
- Buckwalter, Wesley and John Turri. "Inability and Obligation in Moral Judgment." *PLoS ONE* 10, no. 8 (2015): 1-20. doi:10.1371/journal.pone.0136589.
- Chituc, Vladimir, Paul Henne, Walter Sinnott-Armstrong, and Felipe De Brigard. "Blame, Not Ability, Impacts Moral 'Ought' Judgments for Impossible Actions: Toward an Empirical Refutation of 'Ought' Implies 'Can.'" *Cognition* 150 (2016): 20-25. Accessed June 5, 2017. doi: 10.1016/j.cognition.2016.01.013.
- Cohen, Yishai. "An Analysis of Recent Empirical Data on 'Ought' Implies 'Can.'" *Philosophia* 46, no. 1 (2017): 57-67. doi: 10.1007/s11406-017-9892-2.
- Hannon, Michael. "Intuitions, Reflective Judgments, and Experimental Philosophy." *Synthese* 195 (2018): 4147-4168. <https://doi.org/10.1007/s11229-017-1412-1>.
- Henne, Paul, Vladimir Chituc, Felipe De Brigard, and Walter Sinnott-Armstrong. "An Empirical Refutation of 'Ought' Implies 'Can.'" *Analysis* 76, no. 3 (2016): 283-290. doi: 10.1093/analys/anw041.
- Kant, Immanuel. *Critique of Pure Reason*. Translated by Norman Kemp Smith. London: Palgrave Macmillan, 2007. Original work published in 1787.
- Kissinger-Knox, Alicia, Patrick Aragon, and Moti Mizrahi. "'Ought Implies Can,' Framing Effects, and 'Empirical Refutations.'" *Philosophia* 46, no. 1 (2018): 165-182. doi: 10.1007/s11406-017-9907-z.
- Kurthy, Miklos, Holly Lawford-Smith, and Paulo Sousa. "Does Ought Imply Can?" *PLoS ONE* 12, no. 4 (2017): 1-24. <https://doi.org/10.1371/journal.pone.0175206>.
- Kurthy, Miklos, and Holly Lawford-Smith. "A Brief Note on the Ambiguity of 'Ought'. Reply to Moti Mizrahi's 'Ought, Can, and Presupposition: An Experimental Study.'" *Method* 4, no. 6 (2015): 244-249.
- Leben, Derek. "In Defense of 'Ought Implies Can.'" In *Oxford Studies in Experimental Philosophy, Volume 2*, edited by Tania Lombrozo, Joshua Knobe, and Shaun Nichols, chap. 6, Oxford: Oxford University Press, 2018. Kindle.
- Mizrahi, Moti. "Ought, Can, and Presupposition: An Experimental Study." *Method* 4, no. 6 (2015): 232-243.
- Mizrahi, Moti. "Ought, Can, and Presupposition: A Reply to Kurthy and Lawford-Smith." *Method* 4, no. 6 (2015): 250-256.

-
- ¹ Immanuel Kant, *Critique of Pure Reason*, trans. Norman Kemp Smith (London: Palgrave Macmillan, 2007; original work published in 1787), 473.
- ² Vladimir Chituc et al., “Blame, Not Ability, Impacts Moral ‘Ought’ Judgments for Impossible Actions: Toward an Empirical Refutation of ‘Ought’ Implies ‘Can,’” *Cognition* 150 (2016): 20-25, accessed June 5, 2017, doi: 10.1016/j.cognition.2016.01.013.
- ³ Chituc et al. used other kinds of vignettes too—e.g., some that didn’t involve a promise to another person—but these are more peripheral to their study; they mostly function as a kind of control on the more interesting examples in which a promise is made then broken. To read all of their vignettes, see Chituc et al., “Blame, Not Ability” pages 21-22 and supplementary material from experiment 3 (available at <http://dx.doi.org/10.1016/j.cognition.2016.01.013>). For their core conclusion that I paraphrase in the abstract above, see page 23 of their publication.
- ⁴ The following is published by the same researchers as Chituc et al., and it details the same study: Paul Henne et al., “An Empirical Refutation of ‘Ought’ Implies ‘Can,’” *Analysis* 76, no. 3 (2016): 283-290. doi: 10.1093/analys/anw041.
- ⁵ Alicia Kissinger-Knox, Patrick Aragon, and Moti Mizrahi, “‘Ought Implies Can,’ Framing Effects, and ‘Empirical Refutations,’” *Philosophia* 46, no. 1 (2018): 165-182.
- ⁶ Moti Mizrahi, “Ought, Can, and Presupposition: An Experimental Study,” *Methodes* 4, no. 6 (2015): 232-243.
- ⁷ Mizrahi, Moti, “Ought, Can, and Presupposition: A Reply to Kurthy and Lawford-Smith,” *Methodes* 4, no. 6 (2015): 250-256.
- ⁸ Wesley Buckwalter and John Turri, “Inability and Obligation in Moral Judgment,” *PLoS ONE* 10, no. 8 (2015): 1-20. doi:10.1371/journal.pone.0136589.
- ⁹ Miklos Kurthy and Holly Lawford-Smith, “A Brief Note on the Ambiguity of ‘Ought’. Reply to Moti Mizrahi’s ‘Ought, Can, and Presupposition: An Experimental Study,’” *Methodes* 4, no. 6 (2015): 244-249.
- ¹⁰ Miklos Kurthy, Holly Lawford-Smith, and Paulo Sousa, “Does Ought Imply Can?” *PLoS ONE* 12, no. 4 (2017): 1-24, <https://doi.org/10.1371/journal.pone.0175206>.
- ¹¹ Derek Leben, “In Defense of ‘Ought Implies Can,’” in *Oxford Studies in Experimental Philosophy, Volume 2*, eds. Tania Lombrozo, Joshua Knobe, and Shaun Nichols (Oxford: Oxford University Press, 2018), chap. 6, Kindle.
- ¹² Michael Hannon, “Intuitions, Reflective Judgments, and Experimental Philosophy,” *Synthese* 195 (2018): 4150-4154, <https://doi.org/10.1007/s11229-017-1412-1>.
- ¹³ Yishai Cohen, “An Analysis of Recent Empirical Data on ‘Ought’ Implies ‘Can,’” *Philosophia* 46, no. 1 (2017): 57-67. doi: 10.1007/s11406-017-9892-2.
- ¹⁴ Chituc et al. fall into the category of relying heavily, but not exclusively, on quantitative methods. For some of their experiments, they include a free-write space for participants to explain their quantitative responses, but the explanations are brief and play almost no role in Chituc et al.’s total analysis. See Chituc et al., “Blame, Not Ability,” 21-22.
- ¹⁵ I got the idea to use the think aloud method in my x-phi study from the following publication: James Andow, “Qualitative Tools and Experimental Philosophy,” *Philosophical Psychology* 29, no. 8 (2016): 1130, accessed December 22, 2016, doi: 10.1080/09515089.2016.1224826.