EPISTEMIC INDEFINITES

Exploring Modality Beyond the Verbal Domain
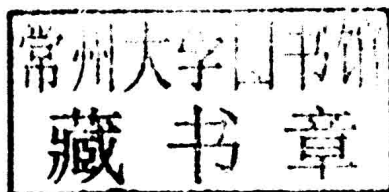
EDITED BY

Luis Alonso-Ovalle &
Paula Menéndez-Benito

Epistemic Indefinites

Exploring Modality Beyond the Verbal Domain

Edited by
LUIS ALONSO-OVALLE AND
PAULA MENÉNDEZ-BENITO

OXFORD
UNIVERSITY PRESS

Epistemic Indefinites

About the contributors

MARIA ALONI is Assistant Professor at the University of Amsterdam. She received her Ph.D. in Philosophy (2001) from the same university. From 2007 to 2012 she coordinated a VIDI research project on indefinites from a cross-linguistic perspective funded by The Netherlands Organisation for Scientific Research (NWO). Her main research interests include quantification, intensionality, questions and imperatives, presupposition and implicature, indefinites, free choice phenomena and negative polarity.

LUIS ALONSO-OVALLE is Associate Professor at McGill University. He received his Ph.D. in Linguistics (2006) from the University of Massachusetts Amherst. His research focuses on natural language semantics and its interfaces with pragmatics. His published work in these areas include articles in *Natural Language Semantics*, the *Journal of Semantics*, and *Linguistics and Philosophy*.

GENNARO CHIERCHIA is Haas Foundations Professor of Linguistics at Harvard University. His books include *Dynamics of Meaning* (Chicago, 1995) and *Logic in Grammar: Polarity, Free Choice, and Intervention* (OUP, 2013). He has also worked on theories of properties and predication, DP structure from a cross-linguistic perspective, the mass–count distinction, generics, and scalar implicatures. He pursues these and related lines of research through both theoretical modelling and experimental methods.

CLEO CONDORAVDI is Professor (Research) at Stanford University. Her areas of research are formal and computational semantics and pragmatics. Topics she works on include modality and temporality, event semantics and the logic of modification, clause types and speech acts, polarity sensitivity and free choice, semantic change, computational models of textual inference, language and reasoning.

ANAMARIA FĂLĂUȘ is a CNRS researcher at the Laboratoire de Linguistique de Nantes, France. She obtained her Ph.D. at the University of Nantes in 2009 and did post-doctoral research at the University of the Basque Country. Her work focuses on formal semantics and its interface with syntax and pragmatics. Recent publications include '(Partially) Free Choice of Alternatives' (*Linguistics and Philosophy*) and an edited volume on alternatives in semantics (Palgrave Macmillan).

STEFAN HINTERWIMMER is lecturer and researcher ('Akademischer Rat auf Zeit') at the Department of German Language and Literature of the University of Cologne. He received his Ph.D. at Humboldt-Universität zu Berlin in 2006. He works on issues in semantics and pragmatics, with a special focus on event/situation quantification, conditionals, specificity and the interaction of information structure and truth-conditional semantics.

TANIA IONIN is an Associate Professor of Linguistics at the University of Illinois at Urbana-Champaign. She received a Ph.D. in Cognitive Science from MIT in 2003. Her research focuses on the interpretation of noun phrases, determiners, and related expressions in both native and

non-native grammars. She has published her work in a variety of journals, including *Bilingualism: Language and Cognition*, *Journal of Semantics*, *Language Acquisition*, *Language Learning*, *Lingua*, *Natural Language Semantics*, and *Second Language Research*. She is an associate editor of the journal *Language Acquisition*.

HSIU-CHEN DAPHNE LIAO got her degree at Harvard University in 2011 with a thesis entitled *Alternatives and Exhaustification: Non-Interrogative Uses of Chinese Wh-Words*. She is currently Assistant Professor at the Department of Languages, Literatures and Linguistics/Graduate Institute of Foreign Literatures and Linguistics at National Chiao Tung University. She has published a number of papers on *wh*-dependencies in Mandarin Chinese and on distributive quantification.

LISA MATTHEWSON is Professor of Linguistics at the University of British Columbia. Her research interests centre on semantic variation and universals, with specific attention to modality, tense, aspect, quantification, and presupposition. She has worked on St'át'imcets (Salish) since the early 1990s, on Gitksan (Tsimshianic) since 2010, and has recently begun investigating Niuean (Austronesian) and Tlingit (Na-Dene).

PAULA MENÉNDEZ-BENITO is a Marie Curie fellow at the Universitat Pompeu Fabra. She received her Ph.D. in Linguistics (2005) from the University of Massachusetts at Amherst. Her research focuses on natural language semantics and its interfaces with pragmatics. Her published work in these areas includes articles in *Natural Language Semantics* and the *Journal of Semantics*.

RICK NOUWEN is Associate Professor at Utrecht University. He received his Ph.D. in Linguistics (2003) from that same university. He has been a post-doctoral affiliate of Edinburgh University, the J.W. Goethe Universität Frankfurt am Main, the Massachusetts Institute of Technology as well as Utrecht University. Since 2013, he has been the principal investigator of a European Research Council 'Consolidator' research programme. He is also the current managing editor of the *Journal of Semantics*. His main research interests include quantification, anaphora, scalarity and degree, expressives, and (multi)dimensionality.

ANGELIKA PORT is a Ph.D. student at the University of Amsterdam. Her research interests include the semantics and pragmatics of (epistemic) indefinites, the use and functions of discourse particles and in general modality. Furthermore, she is interested in diachronic studies and the development of certain expressions over time.

KYLE RAWLINS is an Assistant Professor in the Cognitive Science department at Johns Hopkins University, with a Ph.D. in linguistics from UC Santa Cruz. His research focuses on the semantics and pragmatics of clauses and clause-like objects, such as interrogatives and free relatives, as well as the grammar of modification.

BENJAMIN SLADE is Assistant Professor of Linguistics at the University of Utah. He completed his dissertation on Q-particles in Sinhala and other languages at the University of Illinois under the direction of Hans Henrich Hock in 2011. His published works treat topics including: the morphosyntax of verb–verb collocations in Indo-Aryan from diachronic and synchronic perspectives, the use of neologisms in online cyberpunk discourse, the rise of do-periphrasis in English, the shared inheritance of dragon-slaying formulae in Beowulf and the Rigveda, the role

of Q-particles in Sinhala and Dravidian relative clauses, and the structure of innovative mor-phological processes in Rastafarian speech. He is co-editor with Shu-Fen Chen of *Grammatica et Verba / Glamor and Verve* (Beech Stave, 2013), a Festschrift for Hans Henrich Hock.

CARLA UMBACH is a researcher at the Zentrum für Allgemeine Sprachwissenschaft (ZAS) in Berlin. Her work is in semantics with a Cognitive Science background including experimental approaches and techniques from Artificial Intelligence. Topics of her papers are, for example, contrast, discourse structure, demonstrative pronouns, evaluative predicates, and expressions of similarity.

Abbreviations

1PL	first person plural
2SG	second person singular
AO & MB	Alonso-Ovalle and Menéndez-Benito
AUX	auxiliary
BFA	background-focus-alternative
C, F & S	Chierchia, Fox, and Spector 2012
CC	conceptual cover
CC-shift	conceptual cover shift
CF	choice function(s)
CI	conventional implicature
CL	clitic
CONT	continuous
D	domain of individuals
DAs	sub-domain alternatives
DAT	dative
DE	downward entailing
DEF	definite
deoFC	deontic free choice
DP	determiner phrase
DW	domain widening
EEH	Ebert, Ebert, and Hinterwimmer 2013
EI	epistemic indefinite
epiU	epistemic unknown
EXIS	existential
FC	free choice
FCI	free choice implicature / free choice item
FR	free relative
FUT	future
i-alternative	identity alternative
IMP	imperfective

INDEF	indefinite
INF	infinitive
ISR	intermediate scope reading
K	kernel
MP	Maximize Presupposition!
MV	modal variation
NEG	negation
NP	noun phrase
NPI	negative polarity item
NSR	narrow scope reading
PFV/PERF	perfective
PL	plural
PRT	particle
PS	polarity system
PSIs	polarity sensitive items
PST	past
QP	quantifier phrase
REFL	reflexive
SA	scalar alternative
SG	singular
SI	scalar implicature
SNPIs	strong negative polarity items
spMV	epistemic modal variation
SU	Specific unknown
SUBJ	subjunctive
TOP	topic
vF & G	von Fintel and Gillies 2010
WNPIs	weak negative polarity items
WSR	wide scope reading

Contents

1

Epistemic indefinites: An overview

LUIS ALONSO-OVALLE AND
PAULA MENÉNDEZ-BENITO

1.1 Introduction

Modal expressions in natural language allow us to talk about possible scenarios and unrealized possibilities. For instance, the sentence in (1a), with the modal auxiliary *might*, describes a possible scenario compatible with the available evidence; the one in (1b), with *must*, evokes possible situations in which John does what his parents want him to do.

(1) a. Given what we know, John might be the murderer.
    b. Given his parents' orders, John must be in bed by ten o'clock.

Modality has been extensively studied, both in linguistics and in philosophy (see Portner 2009 and Hacquard 2011 for recent overviews.) Until recently, the study of modality had focused mostly on verbal expressions: modal auxiliaries like *must* or *might*, and attitude verbs like *want* or *believe*. However, the expression of modal notions is not limited to the verbal domain (see, e.g., Kratzer 1981). The sentences in (2) provide some examples of nominal expressions that convey modality.

(2) a. John bought at least four books.
    b. Whatever book John bought was very expensive.
    c. John's books were too expensive.
    d. John bought the wrong books.
    e. John is the man to fix the sink.
    f. The price if you pay now is reasonable.

In (2a), the modified numeral *at least four* triggers the inference that the speaker is uncertain about how many books John bought—he *might* have bought four, or he *might* have bought more. In (2b), *whatever* also conveys speaker's ignorance—the speaker does not know what book John bought. In (2c), *too* adds the implication that John's books *should have been* less expensive. The use of the adjective *wrong* in (2d)

indicates that John *should not* have bought the books that he actually bought. The infinitival relative clause in (2e) conveys that John is *supposed to* fix the sink. Finally, the adnominal conditional in (2f) evokes possible scenarios in which you pay now.

Over the last few years, a body of work on non-verbal modality has emerged. In the recent literature, we find work on modified numerals like *at least n* or *at most n* (see, e.g., Geurts and Nouwen 2007; Büring 2008; Nouwen 2010a; Schwarz 2011a; Coppock and Brochhagen 2013), free relatives of the *whatever*-type (for instance, Dayal 1997; von Fintel 2000; Condoravdi 2005; Tredinnick 2005; Rawlins 2008; Heller and Wolter 2011), *too* constructions (Heim 2000; Hacquard 2000; Meier 2003), infinitival relative clauses as in *the man to fix the sink* (Bhatt 1999; Hackl and Nissembaum 2012), modal adjectives (e.g. Abusch and Rooth 1997; Larson 2000; Schwarz 2006), adnominal conditionals of the form *the price if you pay now* (Lasersohn 1996; Frana forthcoming), and modal indefinites (see references in the next section).

A substantial part of the research on modal indefinites focuses on epistemic indefinites. These are indefinite determiners or indefinite pronouns that signal ignorance on the part of the speaker, thereby conveying information about her epistemic state.[1] German *irgendein* and Spanish *algún* are two cases in point. By using *irgendein* in (3a) and *algún* in (3b), the speaker signals that she cannot identify the doctor that María married. Hence, it would be odd to add a *namely* continuation that explicitly identifies the individual in question, as in (4). Likewise, it would be pragmatically inadequate for the addressee to ask who the doctor is.

(3) a. Maria hat irgendeinen Arzt geheiratet.
        Maria has IRGENDEINEN doctor married
        'Maria married some doctor or other.'

     b. María se casó   con algún   médico.
        María SE married with ALGÚN doctor
        'María married some doctor or other.'

(4) a. Maria hat irgendeinen Arzt geheiratet (#und zwar   Dr. Smith.)
        Maria has IRGENDEINEN doctor married   and indeed Dr. Smith
        'Maria married some doctor or other, namely Dr. Smith.'

     b. María se casó   con algún   médico (#en concreto con el Dr.
        María SE married with ALGÚN doctor,   in particular with the Dr.
        Smith.)
        Smith
        'María married some doctor or other, namely Dr. Smith.'

---

[1] Following the use in Alonso-Ovalle and Menéndez-Benito (2003) and adopted, e.g., in Aloni and Port (2013), we will take the term 'epistemic indefinites' to exclude indefinites that convey speaker's knowledge, such as English *a certain*, German *ein bestimmt*, or French *un certain* (see, e.g., Hintikka 1986; Kratzer 1998, Farkas 2002b; Jayez and Tovena 2006; Hinterwimmer et al. 2013; Martin 2013; among many others). In contrast, Jayez and Tovena (2006) use 'epistemic determiners' as a general term covering both ignorance and knowledge items.

Although the analysis of indefinite phrases has been at the forefront of research in formal semantics for the last three decades, epistemic indefinites did not receive attention until quite recently. In the early 1980s, the study of indefinites motivated the development of radically new theories of natural language interpretation (Kamp 1981; Heim 1982). At that time, and over the following two decades, the literature on the topic focused mostly on the anaphoric and scopal properties that set indefinites apart from other natural language quantifiers. (See, among others, Farkas 1981; Fodor and Sag 1982; King 1988; Ruys 1992; Abusch 1994; Cresti 1995; Reinhart 1997; Winter 1997; Kratzer 1998; Matthewson 1999).

In 1974, however, Strawson had already called attention to another puzzling property displayed by some indefinites, by briefly discussing the behaviour of English singular *some*. Strawson (1974) considers, among others, the examples in (5) and suggests

that the choice of 'some' rather than 'a' embodies what might be called an acknowledgement or recognition of the fact that the identification supplied, though perhaps the best the speaker can do, might be regarded as inadequate to the circumstances of the case; and that the kind of identification which the choice of 'some' rather than 'a' indicates or suggests inability to provide (though perhaps sometimes accompanied by indifference to or unconcern about) may be either further kind-identification or individual identification. (Strawson 1974: 92)

(5) a. Some general has been shot.
     b. Some cabinet minister has been shot.
     c. Some V.I.P. has been shot.   (Strawson 1974: 92)

Despite Strawson's remarks, epistemic indefinites remained largely ignored in the semantic literature until the late 1990s. In 1997, Martin Haspelmath published an influential typological survey (Haspelmath 1997) that included '(lack of) knowledge of the speaker' as one of the possible dimensions of variation within the class of indefinites. He noted that this dimension had not received much attention in the literature, even though it is cross-linguistically widespread:

the semantic distinction of (lack of) knowledge of the speaker is not nearly as important as the two factors discussed in the previous sections, negative polarity and (non-) specificity. It has received very little attention in the theoretical literature. However, it can be found in quite a few languages. In my 40-language sample, I found evidence for such a distinction in ten languages. (Haspelmath 1997: 46)

Since 1997, a substantial body of work on epistemic indefinites has emerged. Some epistemic indefinites that have been discussed in the semantics literature are English singular *some* (Becker 1999; Farkas 2002b; Weir 2012), German *irgendein* (Kratzer and Shimoyama 2002; Aloni 2007a; Lauer 2010; Port 2010; Aloni and Port 2013,[2] Chierchia

---

[2] While Aloni and Port's paper was published in 2013, it has been available on-line since 2010.

2013a), Spanish *algún* (Alonso-Ovalle and Menéndez-Benito 2003, 2008, 2010, 2011a, Giannakidou and Quer 2013), Catalan *algun* (Giannakidou and Quer 2013), Greek *kapjos* (Giannakidou and Quer 2013), French *quelque* (Jayez and Tovena 2007, 2013), *un NP quelconque* (Jayez and Tovena 2002, 2006) and the *n'importe qu-* series (Jayez and Tovena 2002, 2006; Zabbal 2004), Italian *(un) qualche* (Zamparelli 2007; Aloni and Port 2013; Chierchia 2013a), Romanian *vreun* (Farkas 2002a, 2006, Fălăuş 2009, 2011a, 2011b, 2012, 2014), the *-to* series in Russian (Yanovich 2005; Geist 2008), the *-nibud* series in Russian (Geist 2008), the *vola-* and *si-* series in Slovak (Richtarcikova 2013), the *-kin* series in Finnish, Malayalam *wh- oo* (Slade 2011), the Japanese *wh- ka* indeterminates (Sudo 2010; Kaneko 2011; Alonso-Ovalle and Shimoyama 2014), and the Czech *-si* indefinites (Šimík forthcoming).

While the research listed in the previous paragraph has considerably improved our understanding of epistemic indefinites, we are still very far from having a full grasp of the phenomenon. This is so for at least two reasons: first, we still lack enough data to develop a semantic typology of these items, since only a few epistemic indefinites in a few languages have been investigated; second, there is currently no agreement as to what the source of the ignorance effect is. On top of this, the parallelisms and differences with other items that convey epistemic modality, both in the nominal and the verbal domain, have not been sufficiently explored. This situation calls for a research agenda that aims to provide an explanatory semantic typology of epistemic indefinites and to place them within a general typology of modal expressions.

The articles collected in this book bring us one step closer to achieving this goal. In this volume, the reader will find novel empirical observations and important theoretical insights on epistemic indefinites, together with discussions of neighbouring topics (indefinites that convey knowledge, modal free relatives, modified numerals, and epistemic modals), which we think will be ultimately crucial for our understanding of how modality is expressed across categories.

The purpose of this introduction is to contextualize the chapters included in this volume by (i) providing a descriptive overview of the topic and (ii) situating the contributions of the individual chapters against the backdrop of research on epistemic indefinites. Section 1.2 surveys the empirical landscape—it presents and illustrates the parameters of variation that have been identified in the literature on epistemic indefinites. Section 1.3 discusses the role that the articles collected here play on some of the core debates concerning epistemic indefinites. Due to space constraints, we will not attempt to do justice to the rich theoretical literature on the subject, but will focus exclusively on the topics that are closely connected to the contributions to this volume.

1.2 The phenomenon

As noted above, only a handful of epistemic indefinites have been investigated in detail. Furthermore, the available studies on epistemic indefinites have been carried

out within different theoretical frameworks, and therefore a direct comparison of their results is not always possible. Despite being fragmentary, however, the research on epistemic indefinites to date has already identified a number of parameters along which these items can vary. Sections 1.2.1 to 1.2.5 briefly present and illustrate these dimensions of variation. We should note that not all the epistemic indefinites discussed in the literature have been described with respect to all these parameters, and, therefore, we will not be able to present a complete classification of these items.

1.2.1 *Types of ignorance*

All epistemic indefinites convey ignorance—they make an existential claim and signal that the speaker (or some other agent[3]) does not know who (or what) the witness of this claim is[4]). But these items can express different types of ignorance: on the one hand, epistemic indefinites may vary with respect to how much knowledge about the witness counts as knowing who the witness is (Section 1.2.1.1); on the other, some epistemic indefinites can impose restrictions on what types of entities (types vs. tokens) the ignorance component targets (Section 1.2.1.2).

1.2.1.1 What counts as not knowing   Alonso-Ovalle and Menéndez-Benito (2003) observe a contrast between English *some* and Spanish *algún* that shows that these items are sensitive to different types of knowledge. A speaker that sees a professor dancing in front of her, and has never seen the professor before, can felicitously utter the sentence in (6). In this context, the speaker cannot name the dancing professor, but there is a sense in which she knows who the professor is—she can point at him. This way of knowing is compatible with the use of *some*. In contrast, (7), with *algún*, is odd in that situation, suggesting that being able to point at the witness is enough to block the use of *algún*.

(6)   Look! Some professor is dancing on the table!

<div align="right">(Alonso-Ovalle and Menéndez-Benito 2003: 4)</div>

---

[3] Many epistemic indefinites can express ignorance on the part of the some individual other than the speaker. See, e.g., Jayez and Tovena (2007) on French *quelque*, Jayez and Tovena (2002) on French *un quelconque*, and Fǎlǎuş (2009, 2014) on Romanian *vreun*. In contrast, Šimík (forthcoming) claims that Czech *-si* indefinites can only convey speaker's ignorance. In the description that follows, we will focus on the default case, where the ignorance component targets the speaker.

[4] Richtarcikova (2013) notes that there are some contexts where Slovak epistemic indefinites are compatible with speaker's knowledge. A speaker who is able to identify the witness can use one of these indefinites (i) to signal that the identity of the witness is irrelevant, (ii) to indicate that the *hearer* is unable to identify the witness, and (iii) to create 'a sense of suspense', after which the speaker proceeds to identify the witness. The irrelevance component—a 'don't care' effect—has also been reported for indefinites such as *irgendein* (e.g. Kratzer and Shimoyama 2002; Aloni 2007b), *un quelconque* (Jayez and Tovena 2006 and references therein), and *some* and *some or other* (Becker 1999; Farkas 2002b).

(7) # Mira! Algún profesor está bailando encima de la   mesa!
     Look! ALGÚN professor is   dancing on   of the table.
     'Look! Some professor is dancing on the table!'

                                  (Alonso-Ovalle and Menéndez-Benito 2003: 4)

Aloni and Port (2013) note that the same contrast holds between *irgendein* (which behaves like *some*) and Italian *un qualche* (which behaves like *algún*), and identify a number of further contrasts involving different methods of identifying the individual that satisfies the existential claim. Richtarcikova (2013) observes that the *-vola* and *-si* indefinites in Slovak pattern with *some* in the examples above.

1.2.1.2 Types vs. tokens   Epistemic indefinites like *some* or *algún* can express both token-ignorance and kind-ignorance. For instance, (8a) says that the speaker cannot identify the kind of plant that is growing through his wall, while (8b) conveys that the speaker cannot identify the individual file that is infected (Weir 2012). The examples in (9a) and (9b) illustrate these two uses for *algún*.

(8) a. There's some plant growing through the wall of my room.   (Weir 2012: 201)
     b. The hackers implanted a virus into some file on this computer.

                                                              (Weir 2012: 196)

(9) a. En la   pared de mi habitación está creciendo alguna   planta.
        In the wall   of my room   is   growing   ALGUNA plant
        'Some plant is growing through the wall of my room.'

     b. Algún   fichero está infectado.
        ALGÚN file   is   infected
        'Some file is infected'

In contrast, Alonso-Ovalle and Shimoyama (2014) discuss a contrast between two epistemic indefinites in Japanese—*dore-ka* ('which-ka') and *nani-ka* ('what-ka')—which differ with respect to the type vs. token dimension. In a context where the speaker can point to the mushroom that she touched but doesn't know which type of mushroom it is, the sentence in (10b) below, with *nani-ka*, is fine, but its counterpart in (10a), with *dore-ka*, is not.

(10) a. Dore-ka   kinoko-ki   sawat-ta!
         which.one-KA mushroom-DAT touch-PAST
         'I touched some mushroom!'

      b. Nani-ka   kinoko-ni   sawat-ta!
         what-KA mushroom-DAT touch-PAST
         'I touched some mushroom!'

                                      (Alonso-Ovalle and Shimoyama 2014: 17)