Tel Aviv University

The Lester and Sally Entin Faculty of Humanities

School of Philosophy, Linguistics and Science Studies

Copulative perception verbs: What ‘looks’ can teach us about evidentiality and synesthesia

Thesis for the degree of

“DOCTOR OF PHILOSOPHY”

by

Alon Fishman

Prepared under the guidance of Mira Ariel and Yeshayahu Shen

June 2023

Table of Contents

Preface 6

1 Hebrew copulative perception verbs 9

1.1 Introduction 9

1.2 Background 12

1.3 Corpus study 20

1.4 Discussion 29

1.5 Experimental study 36

1.6 Conclusions 39

2 Evidential uncertainty 43

2.1 Introduction 43

2.2 Background 46

2.3 Experimental studies 51

2.4 Computational model 60

2.5 Conclusions 68

3 Synesthetic metaphors 69

3.1 Introduction 69

3.2 Experimental study 76

3.3 Discussion 84

3.4 Conclusions 86

4 General discussion 91

4.1 Summary 91

4.2 Contributions 94

4.3 Future work 96

References 99

Hebrew summary 109

Acknowledgements

I owe thanks to many people, without whom this dissertation would never have come to be. I only have room to thank a few of them here.

First and foremost, to my extraordinary advisors. Mira Ariel is the platonic ideal of a thesis advisor: always available and dependable, equally encouraging and demanding, and constantly on the lookout for both linguistic examples and career opportunities. Each file I sent for review would be back in the blink of an eye, accompanied by high praise even as the text itself would be brimming with well-deserved corrections and criticism. Thank you for opening my eyes to what linguistics can and should be.

Yeshayahu Shen would never let me call him a linguist, but I hope he will accept me calling him a deep thinker, an intent listener, and a mentor. I have come away from every one of our conversations not only knowing something new, but also thinking in a new way about something I'd thought I already knew. Thank you for expanding my horizons beyond linguistics to the wider world of cognitive science.

Other researchers have greatly contributed to this dissertation, directly or indirectly. For their thorough and insightful evaluations, I am indebted to Klaus von Heusinger, Bodo Winter, and Francesca Strik-Lievers. Thank you for making this dissertation much better than it would've been. Judith Degen was not an advisor on this dissertation, but sometimes that was hard to remember. No one else has taught me quite so much in quite so little time. Thank you for hearing out a half-baked idea and helping me turn it into a valuable research project.

I could not have written this dissertation without first becoming enamored with linguistics, thanks to the excellent professors at Tel Aviv University. Special mentions go to Evan-Gary Cohen, who first made me want to grow up to be a linguist, to Nirit Kadmon, who taught me to take credit and responsibility for my original claims, to Outi Bat-El, who drilled into me that a good paper is a finished paper, and to Fred Landman, who taught me that informal rigor is just as important as formal rigor.

I was fortunate to make good friends on my very first semester studying linguistics. And when they all left one year later, I was fortunate to make even better friends, who stuck around longer. To Hila Davidovitch, Daniel Asherov, and Noa Geller, my domestic partners at both the linguistics

department and the linguistics apartment, thank you for talking shop, and for getting sick of talking shop and talking about anything else. To Nicole Katzir and Yechezkel Shabanov, my fellow pragmaticists, thank you for useful discussions, useless ramblings, and usage-based jokes. For pointing me towards the glorious example in the preface, I am indebted to Yael Firer.

The only way I could have studied linguistics in the first place was through the support of my parents, Liora and Ami, who graciously gave up the dream of another MD in the family. Thank you for trying to read each one of my papers.

Writing a dissertation can be grueling, especially through a global pandemic, emigration to the other side of the world, and loss of loved ones. It helps to have someone you love, to proofread all your papers, listen to your endless tangents, commiserate when you grumble, push back when you grumble too much, and remind you at the end of every day of all your good fortune. Thank you, Brightness.

List of symbols and abbreviations

¬	negation
∈	element of
×	cartesian product
∝	proportional to
∑	sum of
$<a_1,\ldots,a_n>$	ordered *n*-tuple
$exp(x)$	*e* raised to the power of *x*
$P(x\|y)$	probability of *x* given *y*
$D_{KL}(P\|\|Q)$	Kullback-Leibler divergence of *P* from *Q*
DIR	evidence directness
F	feminine
M	masculine
PL	plural
REL	evidence reliability
SG	singular
CPV	copulative perception verb
DP	determiner phrase
NP	noun phrase
RSA	Rational Speech Act
[g]	retrieved from Google
collostr.	collostructional
exp.	expected
lit.	literal gloss
obs.	observed

Preface

In an interview on the talk show Late Night with Seth Meyers, actress Rachel Bloom described an online comment on a video of hers (Late Night with Seth Meyers, 2016; emphases mine):

> Some guy wrote on a video I was in "**she looks like her butt stinks**". So here's the thing. I can't… So it's like… This is like five years later, I can't… I don't know if that means "**she looks like she has a stinky butt**", which I do, as does everybody. Uh, or, if it's like "she LOOKS like her butt STINKS", like "**she looks as bad as her butt stinks**".

The comment quoted by Bloom demonstrates the breadth and flexibility of the linguistic domain of perception. On one hand, sensory perception is the main way by which we obtain information about the world. We can therefore use expressions associated with perception to communicate to each other *how* we've come to believe what we believe. For example, the writer of the quoted comment could have been communicating that visual evidence is what led him to the belief that Bloom "has a stinky butt".

On the other hand, we are often interested in percepts: appearances, sounds, odors, flavors, and sensations, as objects in and of themselves. We can use a myriad of expressions to talk about percepts and describe their objective and subjective properties. We can even "borrow" expressions associated with one sensory modality to describe a percept of a different sensory modality. For example, the writer of the comment quoted by Bloom could have been describing Bloom's appearance by likening it to an odor.

In this thesis, I approach the domain of perception both as an object of research, and as a means of exploring other linguistic phenomena. I focus on an understudied class of verbs, here called copulative perception verbs, which includes English *look* and *sound* and Hebrew *nire* 'look' and *nishma* 'sound'. In addition to advancing our understanding of this class of verbs, I mobilize its unique properties to tackle ongoing debates in two domains intersecting perception: evidentiality – the communication of one's source of information; and synesthetic metaphor – the projection of properties from one sensory modality to another.

The first aim of this thesis is to expand the language coverage and strengthen the empirical footing of what we know about copulative perception verbs, as well as provide new insights into the class. I accomplish this by conducting an in-depth quantitative investigation of these verbs in Hebrew,

where they exhibit a grammatical alternation which has not been discussed in previous research (Gisborne 2010, Landau 2011, Poortvliet 2018, Muñoz 2019, inter alia). This investigation is the topic of chapter 1.

The second aim of this thesis is to address an open question in the study of evidentiality, concerning the inference of uncertainty which evidential expressions typically give rise to (Pogue & Tanenhaus 2018, Degen et al. 2019). This inference has previously been attributed to either extralinguistic reasoning about evidence types, or to pragmatic reasoning about alternative utterances (von Fintel & Gillies 2010, Degen et al. 2015). I argue that earlier studies have been unable to settle this debate due to a focus on evidential expressions which imply indirect evidence for a belief. By using copulative perception verbs, I test cases where either direct or indirect evidence is implied. My findings motivate a novel, hybrid account of the uncertainty inference, which involves both extralinguistic and pragmatic reasoning. The debate surrounding the uncertainty inference, my experiments testing it, and a computational implementation of my hybrid account, are covered in chapter 2.

The third aim of this thesis is to address a longstanding issue in the study of synesthetic metaphors, concerning preferences for certain directions over others in metaphorical projection between sensory modalities (Ullmann 1957, Shen 1997, Strik Lievers 2015, inter alia). Theorists have proposed that these directional preferences are caused by either lexical or perceptual factors, but only the former have been established empirically, due to operational difficulties in directly testing or isolating the latter (Petersen et al. 2007, Winter 2016, Strik Lievers & Winter 2018). By using copulative perception verbs, I find the first empirical evidence for a direct effect of perceptual factors on directional preferences, while controlling for lexical factors. The issue of directional preferences in synesthetic metaphors, and my empirical investigation of it, are covered in chapter 3.

In pursuing the aims of this thesis, I draw on a wide range of linguistic frameworks and tools. Theoretically, I engage with and incorporate ideas from formal semantics and pragmatics, usage-based linguistics, and cognitive linguistics. Methodologically, I employ corpus analyses, experimental studies, and computational modelling. Thus, an ancillary aim of the thesis is to show that different approaches to linguistic research can and should be fruitfully combined.

The three chapters of this thesis each stand alone as a self-contained study. Reading and understanding one chapter does not require reading the chapters preceding it. That said, chapter 1 includes an overview of copulative perception verbs, the constructions they occur in, and the meanings they express, which the following chapters refer to.

1 Hebrew copulative perception verbs

1.1 Introduction

The semantic domain of perception is quickly becoming one of the better-studied fields of the lexicon, both within individual languages (e.g., Huumo 2010, Winter 2016) and cross-linguistically (e.g., San Roque et al. 2015, Majid et al. 2018). In addition to the interest it generates in and of itself, research into the lexicon of perception ties into central linguistic concepts such as evidentiality (e.g., Whitt 2010) and subjectivity (e.g., Kaiser 2018).

Verbs of perception have received their fair share of attention in this literature (e.g., Viberg 1983, Ibarretxe-Antuñano 1999, Evans & Wilkins 2000, Gisborne 2010). Since Viberg's (1983) seminal work, a fundamental distinction has been drawn between experiencer-based and phenomenon-based verbs. Experiencer-based verbs take an experiencer as their grammatical subject, and refer to an act of perceiving by that experiencer, whether volitional (e.g., *look (at)* and *listen*) or non-volitional (e.g., *hear* and *see*). Phenomenon-based verbs, in contrast, take a stimulus as their grammatical subject, and refer to a perceptual impression of that stimulus. They may be further divided into copulative verbs, which require a predicate or clausal complement (e.g., *look (like)* and *sound*), and verbs which are predicates in themselves (e.g., *glow* and *buzz*).[1]

There is some overlap between the verb classes described above, in that a single lexical form can belong to more than one class. For example, English *smell* occurs as an experiencer-based verb (e.g., *I smelled it*), a copulative verb (e.g., *It smelled bad*), and a phenomenon-based predicate (e.g., *It smelled*). The different classes may nonetheless be distinguished in such cases, based on their grammatical properties (inflections, argument structure, etc.).

My interest here is in the class of copulative perception verbs (henceforth CPVs), which are often overshadowed in the literature by the other classes of perception verbs. A first clue that these verbs are understudied is the lack of a single accepted label for them. The same class has been referred to as flip perception verbs (Rogers 1974), source-based perception verbs (Viberg 1983), stimulus subject verbs (Levin 1993), physical perception verbs (Asudeh 2002), perceptual resemblance verbs (Asudeh 2004), object-oriented perception verbs (Whitt 2009), *sound*-class verbs (Gisborne

[1] Some phenomenon-based predicates require a complement as well, either an experiencer, e.g., *dazzle*, or a second stimulus, e.g., *reflect*.

2010), perceptual source verbs (Landau 2011), psych predicates (Anand & Korotkova 2018), descriptive perception verbs (Poortvliet 2018), and sensory copulas (Viberg 2019). In addition, some studies simply refer to a construction in which these verbs occur, e.g., *looks like* (Hansen & Markman 2005, Kurumada et al. 2014, McNally & Stojanovic 2017, Kaiser 2018). I adopt the label "copulative" from Taniguchi (1997) because it highlights the unique properties of the class, namely occurrence in a copula position, and taking (some of) the complements of a copula. My use of this term is not intended to endorse an analysis of these verbs as copulas or as copular verbs.[2]

There has been very little work on CPVs in non-European languages. I therefore explore the class of CPVs in Hebrew, focusing on three constructions in which they occur, exemplified in (1-2) below. In (1a) and (2a), the verb takes an adjectival complement that agrees in gender (and number) with the verb's subject. In this sense, the verb behaves like a copula; cf. (3a). I call this the COPULATIVE construction.[3] In contrast, in (1b) and (2b), the verb's subject and complement do not agree, and in fact, the complements in these cases are not adjectival at all, but adverbial (see also Avineri 2021).[4] That is, the verbs in these cases pattern with run-of-the-mill verbs modified by adverbs; cf. (3b). To the best of my knowledge, Hebrew is the first language reported to allow both adjectival and adverbial complements for CPVs.[5,6] I call this the VERBAL construction. Finally, in (1c) and (2c) the verb occurs in an IMPERSONAL construction, either with the expletive subject *ze*

[2] The terms "copulative" and "copular" are sometimes read as interchangeable, particularly in works rejecting both terms' appropriateness for the verbs under discussion (e.g., Aikhenvald & Storch 2013, Staniewski & Gołębiowski 2021). In my reading of the literature, the two terms are distinct: "copulative" is descriptive, referring to copula-like behavior or properties (e.g., Viberg 1983, Taniguchi 1997), whereas "copular" is analytic, referring to extant formal analyses of copulas (e.g., Landau 2011, Poortvliet 2018).

[3] Here too, the term "copulative" is meant to evoke the copula-like nature of the construction. To clarify the relationship between copulative verbs and the copulative construction: the characteristic feature of the former is that they may occur in the latter (along with several other constructions, some of which may be language-specific), though I don't consider this a necessary feature in the strict sense; see fn. 8.

[4] Most Hebrew adverbs are formally identical to the masculine singular forms of the corresponding adjectives, as with *muzar* 'weird.MSG'/'weirdly' in (1b). An exception to the general rule is *axeret* 'different.FSG'/'differently', which is identical to the feminine singular form of the adjective, as in (2b).

[5] Russian CPVs, like Hebrew CPVs, allow both adjectival and adverbial complements, e.g., *viglyadit ploxim* 'looks bad' versus *viglyadit ploxo* 'looks badly' (Daniel Asherov, Avital Fishman, and Avital Zaruvimsky, p.c.). I am not aware of academic work that touches on this alternation.

[6] As a complement of CPVs, English *well* occurs as an adjective, analogous to its use with a copula (contrast *look well*, *be well* with *behave well*). Viberg (2019) observes that the complement of French *avoir l'air* 'look' (lit. 'have the look') can optionally agree with the subject rather than with the noun *air*, but in either case it is adjectival, not adverbial.

or with no subject. Here and throughout the thesis, examples marked with [g] were retrieved from Google.

(1) a. **ani nishma'at muzara** kshe-ani menasa lesaper mashehu. [g] **(**Copulative)
I sound.FSG weird.FSG when-I try.FSG to.tell something
'I sound weird whenever I try to say something.'

b. kshe-ani shara be-ivrit **ani nishma'at muzar**. [g] (Verbal)
when-I sing.FSG in-Hebrew I sound.FSG weirdly
'When I sing in Hebrew I sound weird.'

c. ulay **ze nishma** **she-ani muzara** she-ani shoelet. [g] (Impersonal)
maybe it sounds.MSG that-I weird.FSG that-I ask.FSG
'Maybe it sounds like I'm weird for asking this.'

(2) a. **hu nire** **axer**, muzar [g] (Copulative)
he looks.MSG different.MSG weird.MSG
'He looks different, strange.'

b. ba-tmuna **hu nire** **axeret** legamrey [g] (Verbal)
in.the-picture he looks.MSG differently completely
'In the picture he looks completely different.'

c. ve-gam me-ha-tguvot shelo **nire** **li** **she-hu axer**. [g] (Impersonal)
and-also from-the-comments his looks.MSG to.me that-he different.MSG
'And also from his comments it looks to me like he's different.'

(3) a. raxel hayta muzara. (Copula)
Rachel was.FSG weird.FSG
'Rachel was weird.'

b. raxel hitnahaga muzar. (Verb)
Rachel behaved.FSG weirdly
'Rachel behaved oddly.'

The copulative and verbal constructions exhibit distributional differences, which I take to be clues for semantic differences. I present a large-scale corpus study showing that the constructions vary

in their co-occurrence with different CPVs and with a dative argument, and attract different complements and subjects. In particular, the copulative construction attracts abstract complements (e.g., *hegyoni* 'logical') and abstract subjects (e.g., *she'ela* 'question'), whereas the verbal construction attracts subjective multidimensional complements (e.g., *tov* 'good') and perceivable subjects (e.g., *ish* 'person').

These findings suggest that verbal and copulative constructions map to previously identified meanings of CPVs. The verbal construction has an ATTRIBUTARY meaning: it attributes a property to a perceptual impression. However, the complement slot of the construction is more restricted than predicted by previous accounts of the attributary meaning (e.g., Gisborne 2010, Petersen & Gamerschlag 2014, Muñoz 2019), motivating a novel account in terms of dimension selection (Sassoon 2013). The copulative construction, as well as the impersonal construction, have a PARENTHETICAL meaning: they take and modify a proposition argument. But the two constructions are not equivalent. I present a preference experiment revealing that the copulative construction encodes an evidential meaning, whereas the impersonal encodes an epistemic modal meaning. Consequently, the three-way formal distinction corresponds to a three-way semantic distinction.

1.2 Background

The literature on CPVs is not a cohesive body of work. In the generative tradition, recent studies on CPVs focus on one particular construction known as Copy Raising, e.g., *Rachel looks like she went to a show*, especially on the relationship between the matrix subject and its "copy", the embedded pronoun (Potsdam & Runner 2001, Asudeh 2002, 2004, Landau 2011, Asudeh & Toivonen 2012, but cf. Muñoz 2019). In contrast, studies drawing on cognitive or functional linguistics focus on relating CPVs to other perception verbs, synchronically and diachronically (Taniguchi 1997, Jackendoff 2007, Whitt 2009, 2010, Gisborne 2010, Fernández Jaén 2015).

Other studies do not investigate CPVs per se. Rather, they use CPVs as diagnostics or in experimental materials in order to investigate other issues. Hansen & Markman (2005) use experimental materials with the phrase *looks like* to explore children's ability to distinguish appearance from reality. Kurumada et al. (2014) use experimental materials with the same phrase to explore the real-time processing of visual context and intonation contours. McNally & Stojanovic (2017) propose a diagnostic for distinguishing between certain adjective classes, based on co-occurrence with the CPVs *look* and *sound*. Kaiser (2018) uses experimental materials with

the CPVs *look*, *smell* and *taste* to explore how sensory modality influences the identification of a subjective attitude holder. Pogue & Tanenhaus (2018) use experimental materials with the CPV *looks* to explore the expression and perception of (un)certainty. Chapters 2 and 3 below are also examples of this practice.

The divergent theoretical perspectives, methodologies and terminology manifest in the literature all lead researchers to overlook previous observations or to, inadvertently, "reinvent the wheel". Rather than discuss previous studies individually, below I review recurring themes and noteworthy points of contention, glossing over differences in terminology and theoretical assumptions. Most of the reviewed literature discusses English or other European languages, including German (Whitt 2009, 2010, 2011, Viberg 2019, Staniewski & Gołębiowski 2021), Swedish (Asudeh & Toivonen 2012, Viberg 2019), Spanish (Fernández Jaén 2015, Albelda Marco & Jansegers 2019), Dutch (Poortvliet 2018), Finnish (Viberg 2019), French (Viberg 2019) and Polish (Staniewski & Gołębiowski 2021), along with two previous works touching on Hebrew (Landau 2011, Avineri 2021).

First, most researchers agree that CPVs have more than one meaning. Specifically, most researchers distinguish between an ATTRIBUTARY meaning (following Gisborne 2010), which attributes a property to (a perceptual impression of) an object, and a PARENTHETICAL meaning, which modifies a proposition. The parenthetical meaning might be specified as an epistemic modal, i.e., relating a proposition to an individual's beliefs, or as an evidential, i.e., relating a proposition to a type or source of evidence, or as both. To illustrate the difference between attributary and parenthetical meanings, contrast the exchange in (4), which concerns an auditory impression of Rachel, with the one in (5), which concerns (belief in or evidence for) the proposition 'Rachel is a good comedian'. This basic semantic distinction underlies my review of the literature in this section, as well as my discussion of the corpus findings in section 1.4.

(4) A: Is Rachel a good singer?
B: Yeah, I've heard her sing and **she sounds really good**. (Attributary)

(5) A: Is Rachel a good comedian?
B: I haven't seen her perform, but **she sounds really good**. (Parenthetical)

The distinct meanings of CPVs were pointed out as early as Rogers (1974). Rogers focuses on one meaning of CPVs, which he analyzes as a belief caused by a perception – a parenthetical meaning

which is both evidential and epistemic. But he notes that CPVs have at least two other meanings: one which doesn't presuppose "actual physical perception" and is "roughly synonymous with… *seem*" (p. 90) – epistemic but not evidential; and another which attributes a property to an object only with regard to the sensory modality expressed by the verb (p. 152) – attributary.

Later researchers independently reached similar conclusions, even if couched in different terms. Lasersohn (1995), discussing *sound*, proposes that it's ambiguous between a perceptual use denoting a two-place relation between individuals and properties – attributary; and a raising verb denoting a one-place predicate of propositions – unspecified parenthetical. Usoniene (2000), discussing *look*, distinguishes "the speaker's tentative appraisal of some propositional content" – epistemic; from "descriptions of appearances… based on sensory data" – attributary (pp. 189-190).

Hansen & Markman (2005), discussing the construction *looks like*, propose that it can refer either to "likely reality" – epistemic; or to "outward appearance", which covers both (true) attribution of properties and (mistaken) belief in a proposition (p. 236) – attributary and evidential, respectively. Whitt (2009, 2011), discussing English *look* and *sound* and German *aussehen* 'look' and *klingen* 'sound', distinguishes evidential uses from uses which "merely indicate a value judgment" (2011: p. 349) or where "a mere auditory attribute… is being described" (2009: p. 1085) – both attributary. Viberg (2019), discussing Swedish *se ut* 'look' and *låta* 'sound' and their translations into four European languages, distinguishes evidential and epistemic uses from ones describing appearance or acoustic properties, expressing value judgments, and making comparisons – all attributary.

Gisborne (2010) and Landau (2011), despite vast theoretical differences, both end up with comparable sets of three meanings for CPVs: two evidential, and a third which combines with its complement to form a complex predicate, taking the subject as an argument – attributary. Muñoz (2019) develops separate analyses for CPVs composing evidence-related predicates – evidential; and composing what he calls hybrid experiential predicates, which apply a property "to the sensory modality in question, as stimulated by the subject" (p. 148) – attributary.

Despite the repeated observations reviewed above, the fact that CPVs have multiple meanings nevertheless goes unnoticed at times, which again speaks to the status of CPVs as understudied. Overlooking this fact can have detrimental consequences when the verbs are employed as linguistic diagnostics or in experimental materials. For example, McNally and Stojanovic's (2017) diagnostic for distinguishing between aesthetic and evaluative adjectives on one hand, and

predicates of personal taste on the other hand, is based on co-occurrence with the CPVs *look* and *sound*. However, they assume that the CPVs always indicate source of information – an evidential meaning; and overlook the attributary meaning entirely. Consequently, the contrasts they report could be attributed to different meanings of the CPVs, rather than to the different adjective classes they are interested in, especially since some of their examples seem to allow both meanings (e.g., *The cake looks good to me*). Another example is Kaiser's (2018) experiment designed to investigate how sensory modality influences the identification of a subjective attitude holder, using materials containing the CPVs *look*, *smell* and *taste*. Since she does not address the fact that the CPVs have multiple meanings, her findings regarding differences between sensory modalities might be mediated by different meanings of the CPVs, or muddled by cases where the CPV could have any one of the two meanings (e.g., *The muffin looked disgusting*).

Next, it is well-established that CPVs can freely take adjectival complements, comparative complements, and clausal complements, with the latter also available in impersonal constructions; see (6a-d).[7,8] Other complements are more limited. For instance, English *look* can take infinitive complements, *look* and *sound* can both take nominal complements, and *taste* and *smell* can both take complements headed by *of*; see (7a-d) (examples (7a-c) are from Usoniene 2000). As I show below, Hebrew has counterpart constructions for (6a-d) and (7d), along with a variant taking an adverbial complement, an alternation which has not been discussed in earlier studies.

(6) a. Rachel looks funny.
    b. Rachel looks like a comedian.
    c. Rachel looks like/as if/as though she went to a show.
    d. It looks like/as if/as though Rachel went to a show.

(7) a. Most of them looked to be students.
    b. Confrontation looked a real possibility.
    c. From the way you describe him he sounds a real idiot.

[7] It has been known at least since Heycock (1994) that clausal complements of CPVs do not necessarily contain a copy of the verb's subject.

[8] Polish CPVs cannot take adjectival complements at all, and instead take adverbial complements (Staniewski & Gołębiowski 2021). I nevertheless consider them CPVs by virtue of their family resemblance to CPVs in other languages.

d. It tastes/smells of garlic.

When it comes to pairing (surface) structures with meanings, there is general agreement that CPVs with an adjectival complement can have both attributary and parenthetical meanings (Rogers 1974, Lasersohn 1995, Gisborne 2010, Whitt 2011, Poortvliet 2018, Muñoz 2019, but cf. Landau 2011), as do CPVs with a comparative complement (Lasersohn 1995, Hansen & Markman 2005, Gisborne 2010, Poortvliet 2018, but cf. Rogers 1974). Thus, (6a) above can either attribute funniness to Rachel's appearance, or modify the proposition 'Rachel is funny'. Likewise, (6b) can either compare Rachel's appearance to those of a (generic or specific) comedian, or modify the proposition 'Rachel is a comedian'. The main goal of the corpus study in section 1.3 is to explore which meanings are possible for Hebrew CPVs with adjectival and with adverbial complements.

A number of studies draw explicit parallels between adjectival complements in the attributary use, and adverbs. Postal (1971) proposes that the adjectives in these cases are irregular *-ly* adverbs. Taniguchi (1997) argues that English CPVs developed out of the use of the phenomenon-based predicates *sound* and *smell* with adverbs (e.g., *The alarm sounded loudly*, *The food smelled unpleasantly*), and Poortvliet (2018) makes a similar argument for Dutch *klinken* 'sound' and *ruiken* 'smell'. In Gisborne's (2010) analysis, the way CPVs compose with their complements in attributary uses parallels the way other verbs compose with adverbs. These parallels naturally raise the hypothesis that Hebrew cases with adverbial complements would have attributary meanings.

As for CPVs with a clausal complement, there is again general agreement that they only have parenthetical meanings (Rogers 1974, Lasersohn 1995, Gisborne 2010, Landau 2011, Whitt 2011, Poortvliet 2018, Viberg 2019, Avineri 2021). But disagreements surface as to whether parenthetical meanings are the same for different structures. Rogers (1974) and Lasersohn (1995) don't correlate different structures with different parenthetical meanings. In Gisborne's (2010), Landau's (2011) and Muñoz's (2019) analyses, parenthetical meanings are always evidential, but impersonal constructions differ from other constructions in that they never specify a source of evidence.[9] Hansen & Markman (2005) go further and imply that *looks like* in impersonal constructions can only be an epistemic modal, whereas in other constructions it can be an evidential

[9] According to Muñoz (2019), a non-expletive subject must be the source of evidence, whereas according to Gisborne (2010) it may or may not be. Landau (2011) agrees with the latter, except for clausal complements with no element coindexed with the CPV's subject, i.e., with no "copy", which obligatorily specify the subject as the source of evidence.

instead. In section 1.5, I report on a preference experiment designed to test whether parenthetical meanings differ between two Hebrew constructions.

In addition to the grammatical categories of CPVs' subjects and complements, some studies also address their conceptual categories. In attributary uses, it is typically taken for granted that the verb's subject denotes a stimulus perceivable through the relevant sensory modality (Lasersohn 1995, Gisborne 2010, Landau 2011, Petersen & Gamerschlag 2014). Viberg (2019) points out an exception, in that English *sound* and its counterparts Swedish *låta*, German *klingen* and Finnish *kuulostaa* can be used to form an evaluative judgment – an attributary use – of linguistic content, even when not perceived via hearing, e.g., in written form. In parenthetical uses, a few researchers take the verb's subject to again denote a perceivable stimulus (Ash & Toivonen 2012, Petersen & Gamerschlag 2014), but many others find cases where it denotes something not perceivable through the relevant sense, e.g., an abstract entity (Rogers 1974, Lasersohn 1995, Usoniene 2000, Gisborne 2010, Landau 2011, Poortvliet 2018, Viberg 2019, Avineri 2021).

Regarding the verb's complement, in attributary uses it is taken to denote a predicate over the relevant sensory modality, which is often a value judgment (Rogers 1974: p. 151, Usoniene 2000, Whitt 2009, 2011, Poortvliet 2018, Albelda Marco & Jansegers 2019, Viberg 2019, Staniewski & Gołębiowski 2021). The status of predicates associated with a single sense, as in *look red* or *sound loud*, is contested. While Gisborne (2010) along with Petersen & Gamerschlag (2014) consider these exemplary attributary uses, Viberg (2019) finds that they stand out for "point(ing) to uncertainty or special conditions" (p. 30), echoing the discussion in Grice & White (1961). In parenthetical uses, Petersen & Gamerschlag (2014) argue that the verb's complement is permissible only if it denotes a property inferable through the relevant sensory modality (see also Rogers 1974: p. 138). For example, the complement of *taste* cannot normally denote a shape. In the corpus study presented below, I explore how two Hebrew constructions differ in which complements and subjects they attract.

In most studied languages, CPVs also optionally take a dative argument, which is analyzed as an experiencer, belief-holder, evidence-holder, or some combination thereof. [10] In cases with no dative, this role is filled pragmatically, by default either by the speaker (Whitt 2009, Landau 2011,

[10] Dutch CPVs are an exception in that they cannot take dative arguments (Poortvliet 2018).

Asudeh & Toivonen 2012, Viberg 2019) or by a generic individual (Usoniene 2000, Jackendoff 2007, Muñoz 2019). Gisborne (2010) posits a difference between parenthetical and attributary meanings, in that the former have an obligatory experiencer role, which is filled pragmatically in the absence of a dative argument, whereas the latter have no experiencer role and never take a dative argument. In the corpus study presented below, I examine whether the different Hebrew constructions vary in co-occurrence with a dative argument.

Finally, and notably, most researchers do not propose an explicit procedure for disambiguating uses of CPVs, and only distinguish them by explicating or paraphrasing their meanings. Hansen & Markman (2005) argue that having such a procedure is critical for experimental studies employing CPVs, and a similar argument could be made for coding naturally occurring data. The few such procedures to have been developed are all quite limited in their applicability.

Lasersohn (1995) offers a diagnostic for distinguishing uses of *sound* when it is followed by *like NP*: the attributary meaning is preserved, but the parenthetical is not, when a second mention of *sounds* is added after the NP. Thus, *John sounds like a frog sounds* can only attribute a description to John's sound. Hansen & Markman (2005) claim that an epistemic meaning is the default for *looks like*, but can be blocked if relevant information is made part of the conversation's common ground. Thus, a non-epistemic meaning can be forced by spelling out knowledge of the verb's subject, e.g., *This sponge looks like a rock*, or by adding a qualification, e.g., *This looks like a rock but that's not what it really is*. Gisborne (2010) similarly proposes using the continuation *but (it) isn't really* to distinguish between different uses of CPVs with an adjectival complement. Parenthetical uses, but not attributary uses, allow this continuation. Thus, *Rachel looks funny, but isn't really* can only be about the proposition 'Rachel is funny' and cannot attribute funniness to Rachel's looks. In section 1.4, I apply Gisborne's diagnostic to Hebrew CPVs and show that it has a straightforward grammatical explanation.

1.2.1 Hebrew CPVs

In Hebrew, the class of CPVs consists of *nire* 'look', *nishma* 'sound', *meriax* 'smell' and *margish* 'feel'.[11] Each of these lexemes has other uses in addition to their CPV use. Most prominently,

[11] Some Hebrew speakers accept *toem* or *nitam* 'taste' as a CPV, as in the naturally-occurring examples in (i-ii) (kindly provided by Yuval Katz, p.c.). The former standardly occurs as an experiencer-based verb, and the latter is non-standard according to most dictionaries, but is listed in one online dictionary

meriax and *margish* are also used as experiencer-based verbs (as discussed in detail by Avineri 2021), similarly to English *smell* and *taste*, while *nire* and *nishma* are also used as the passive forms of the experiencer-based verbs *roe* 'see' and *shomea* 'hear', respectively.[12]

Hebrew CPVs can take as their complement an adjectival phrase, an adverbial phrase, a comparative construction headed by *kmo* 'like' or *ke-* 'as', or a clausal complement with the complementizer *keilu* 'as if'. Impersonal constructions – without a subject or with the expletive subject *ze* – additionally allow the complementizers *she-* 'that' and *ki* 'that'. The CPV *meriax* 'smell' can also take a prepositional complement in the form *me-DP* 'of DP'. The various complements are demonstrated in (8a-f). Regardless of which type of complement they take, Hebrew CPVs may optionally take a dative argument.

(8) a. raxel niret /nishma'at /merixa /margisha muzara. (Copulative)
Rachel looks. /sounds. /smells. /feels.FSG weird.FSG
'Rachel looks/sounds/smells/feels weird.'

b. raxel niret /nishma'at /merixa /margisha muzar. (Verbal)
Rachel looks. /sounds. /smells. /feels.FSG weirdly
'Rachel looks/sounds/smells/feels weird.'

C. raxel niret /nishma'at /merixa /margisha kmo komikait.
Rachel looks. /sounds. /smells. /feels.FSG like comedian.F

(www.milog.co.il/נטעם). Occurrences of either form as a copulative verb are rare, and may be explicitly called out as neologisms by interlocutors.

(i) lama kafe be-tvila **toem** axeret me-kafe be-xilxul?
why coffee in-immersion tastes differently than-coffee in-percolation
'Why does immersion-brewed coffee taste different than drip-brewed coffee?'

(ii) nire, nishma ve-meniax she-gam **nitam** nifla!
looks, sounds and-suppose that-also tastes wonderful
'Looks, sounds and probably also tastes wonderful!'

[12] *nire* is additionally used in an adjectival form meaning 'visible', particularly in the phrases *bilti nire* 'invisible' and *nire la-ain* 'visible to the eye', as well as in idiomatic phrases with various modal meanings: *ka-nire* 'probably' (lit. 'as it looks'), *ke-xol ha-nire* 'apparently' (lit. 'to the extent it is seen'), and *nire le-X* 'to X's liking' or 'to X's mind' (lit. 'looks to X'). *nishma* is additionally used to mean 'obey' and in the idiomatic greeting *ma nishma?* 'how are you?" (lit. 'what is heard'). *meriax* is additionally used as a phenomenon-based predicate, similarly to English *smell*, as in the idiomatic phrase *lo meriax lo masriax* 'inoffensive, unremarkable' (lit. 'doesn't smell, doesn't stink').

'Rachel looks/sounds/smells/feels like a comedian.'

d. raxel niret /nishma'at /merixa /margisha keilu hi hayta be-hofa'a.
Rachel looks. /sounds. /smells. /feels.FSG as.if she was.FSG in-show
'Rachel looks/sounds/smells/feels like she went to a show.'

e. (ze) nire /nishma /meriax /margish she- /keilu raxel hayta be-hofa'a.
(it) looks. /sounds. /smells. /feels.MSG that- /as.if Rachel was.FSG in-show
'It looks/sounds/smells/feels like Rachel went to a show.'

(Impersonal)

f. raxel merixa me-sigaryot.
Rachel smells of-cigarettes
'Rachel smells of cigarettes.'

1.3 Corpus study

In this section I present distributional data on Hebrew CPVs extracted from the Hebrew Web 2014 corpus (heTenTen14, with automated part-of-speech tagging developed by Meni Adler) available on www.sketchengine.eu (Kilgarriff et al. 2014). This is a multiple domain corpus of 890 million words, made up of texts collected from the Internet, including newspaper materials, Wikipedia articles, blog posts, and both personal and commercial web pages. The corpus data and the scripts used to analyze them are available online at https://osf.io/g7kw6/.

In presenting the data, I focus on differences between the copulative and verbal constructions, examining their distributions with each CPV, their co-occurrence with a dative argument, and the subjects and complements they attract. The motivation for this focus is predominantly methodological: because the copulative and verbal constructions share an argument structure, namely a subject slot and a (predicate) complement slot, they allow for direct quantitative comparisons. In contrast, the impersonal construction by definition has no subject slot, and its complement is not a predicate but a clause.

To make it possible to distinguish between copulative and verbal constructions, data on them are restricted to feminine and plural inflections of the CPVs.[13] This is because adjectival and adverbial

[13] Hebrew has two categories for grammatical gender (masculine and feminine) and two for number (singular and plural), obligatorily marked on nouns, verbs and adjectives. A few nouns, mostly ones

complements in masculine singular form are indistinguishable, with very few exceptions. To eliminate false positives where *margish* 'feel' occurs as an experiencer-based verb rather than a CPV, data on it are restricted to instances with a dative argument. This is because *margish* takes the same types of complements whether it occurs as a CPV or as an experiencer-based verb, and the latter are far more frequent.[14]

The total occurrences of each CPV in copulative, verbal and impersonal constructions, and their co-occurrence with a dative argument in each construction, are presented in Table 1. Overall, *nire* 'look' is more frequent than *nishma* 'sound' by an order of magnitude, as is *nishma* in comparison to *meriax* 'smell'. The frequency of *margish* 'feel' falls between *nishma* and *meriax*.

Table 1. Total occurrences of each CPV in copulative, verbal and impersonal constructions, and their co-occurrence with a dative argument in each construction. Data on copulative and verbal constructions are restricted to feminine and plural inflections of the CPV. Data on *margish* are restricted to instances with a dative argument.

CPV	*nire* 'look'	*nishma* 'sound'	*meriax* 'smell'	*margish le-* 'feel to'
Copulative	29,281	5,026	9	217
↳ with dative	(6,561)	(723)	(4)	(217)
Verbal	8,616	1,684	217	54
↳ with dative	(219)	(69)	(8)	(54)
Impersonal	53,406	1,357	5	83
↳ with dative	(14,781)	(140)	(0)	(83)

denoting units of time, additionally have a dual form (e.g., *yom* 'day' – *yomaim* 'two days' – *yamim* 'days', *xodesh* 'month' – *xodshaim* 'two months' – *xodashim* 'months'), which agrees with plural verbs and adjectives exactly like the plural form, and is tagged as plural in the Hebrew Web 2014 corpus.

[14] These restrictions were adopted following a preliminary process of extracting and manually coding a random sample of 1000 occurrences of each of the lexemes *nire*, *nishma*, *meriax* and *margish*. In these samples, 74.7% of adjectival and adverbial complements were indistinguishable, and instances of *margish* as a CPV amounted to only 5.5% of its total occurrences. The hand-coded samples are available online at https://osf.io/g7kw6/.

The distribution of each CPV's occurrences in copulative and verbal constructions, with and without a dative argument, are visualized in Figure 1. Copulative constructions are considerably more frequent than verbal constructions for *nire* 'look' (~3:1), *nishma* 'sound' (~3:1) and *margish* 'feel' (~4:1). The opposite is true for *meriax* 'smell' (~1:24). A series of Fisher's exact tests, with Bonferroni corrections for multiple comparisons, reveals that the distribution of constructions differs significantly between *meriax* and each of *nire* ($p < .001$), *nishma* ($p < .001$), and *margish* ($p < .001$).

Across verbs, co-occurrence with a dative argument is considerably higher in the copulative construction (between 14.4% and 44.4%) than in the verbal construction (between 2.5% and 4.1%). A series of Fisher's exact tests, with Bonferroni corrections for multiple comparisons, reveals that co-occurrence with a dative argument differs significantly between copulative and verbal constructions for each of *nire* 'look' ($p < .001$), *nishma* 'sound' ($p < .001$), and *meriax* 'smell' ($p = .002$).

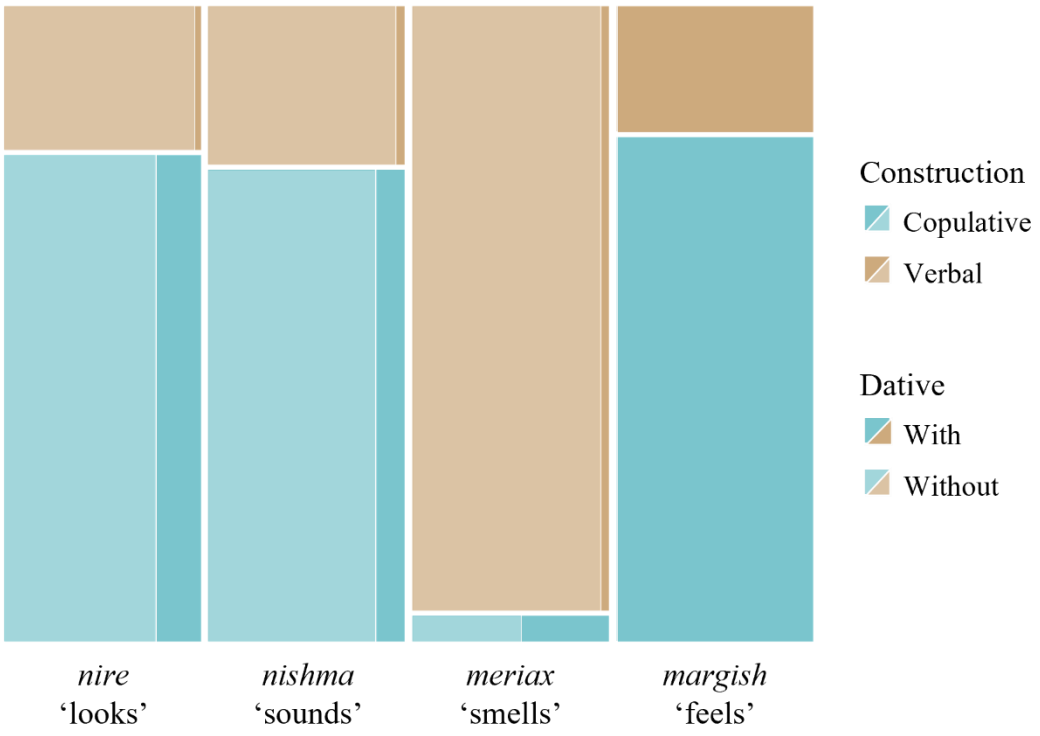


Figure 1. Distribution of CPVs in copulative and verbal constructions. The darker shading represents instances with a dative argument.

Impersonal constructions are most frequent for *nire* 'look' (58.5%), followed by *margish* 'feel' (23.4%) and *nishma* 'sound' (16.8%), and finally *meriax* 'smell' (2.2%). Co-occurrence with a dative argument in the impersonal construction is highest for *nire* (27.7%), followed by *nishma* (10.3%) and finally *meriax* (0.0%).

Next, I employ the exploratory technique of Distinctive Collexeme Analysis (Gries & Stefanowitsch 2004) to identify further differences between copulative and verbal constructions. Distinctive Collexeme Analysis is a method of quantifying the attraction between a lexeme and a construction, in contrast to another construction. Unlike traditional collocational analyses, it focuses on lexemes occurring in particular slots within a construction, rather than all words within a given span.

Here I conduct Distinctive Collexeme Analyses on two slots in the copulative and verbal constructions: the complement and the subject. Due to the vast differences in frequency, I conduct separate analyses for each of the four CPVs. The analyses are performed using Gries' R script (Gries 2014). In each case, the results indicate which lexemes are attracted to the relevant slot in each of the two constructions, and to what extent, expressed in a measure termed collostructional strength.[15]

The complete set of results of all eight Distinctive Collexeme Analyses, comprising a total of 6495 lexeme/verb combinations, is available online at https://osf.io/g7kw6/. I present side by side comparisons of the top 5 complements (Tables 2-5) and the top 5 subjects (Tables 6-9) attracted to each of the two constructions, for each of the four CPVs. These comparisons reveal striking differences between the copulative and verbal constructions, consistent across the four verbs. Each table shows observed and expected frequencies with both constructions, as well as collostructional strength with the preferred construction. For example, the top row of Table 2 shows that the complements of *nire* 'look' most strongly attracted to the copulative and verbal constructions are *hegyoni* 'logical' and *nehedar* 'terrific', respectively. The observed frequencies of *hegyoni* are 566 in the copulative construction and 9 in the verbal construction, whereas its expected frequencies (i.e., if it occurred in either construction randomly) are 444.27 and 130.73, respectively. Its

[15] The collostructional strength of a lexeme with its preferred construction is equal to the negative base-10 logarithm of the *p*-value of a one-tailed Fisher-Yates exact test, performed on the contingency table of that lexeme versus all other lexemes in either of the two constructions.

collostructional strength with the copulative construction is 50.41 (representing a *p*-value under 10^{-50}).

Table 2. Top 5 complements of *nire* 'look' in copulative and verbal constructions.

Preference for copulative construction				**Preference for verbal construction**			
Complement	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength	Complement	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength
hegyoni 'logical'	566 (444.27)	9 (130.73)	50.41	*nehedar* 'terrific'	27 (462.04)	571 (135.96)	∞
mat'im 'fitting'	493 (384.01)	4 (112.99)	48.78	*tov* 'good'	324 (2940.69)	3,482 (865.31)	∞
xaser 'lacking'	471 (366.23)	3 (107.77)	47.80	*axer(et)* 'different'	273 (1355.22)	1,481 (398.78)	∞
pashut 'simple'	481 (374.73)	4 (110.27)	47.46	*metsuyan* 'excellent'	53 (549.35)	658 (161.65)	∞
raxok 'distant'	460 (360.05)	6 (105.95)	42.56	*nifla* 'wonderful'	23 (413.37)	512 (121.63)	296.90

As Table 2 shows, complements of *nire* 'look' attracted to the verbal construction are all subjective in some way (see section 1.4), and are also virtually unrestricted in terms of what they can be predicated of; any object can be terrific, good, different, etc., whether it's concrete or abstract, animate or inanimate, perceivable or non-perceivable. On the other hand, complements attracted to the copulative construction are not necessarily subjective, and some can only be predicated of specific classes of objects, e.g., *hegyoni* 'logical' for abstract ideas and *raxok* 'far' for physical locations. Complements that can only be predicated of specific classes of objects are shaded in gray in Tables 2-5.

In addition, Table 2 shows that the observed frequencies of complements attracted to the verbal construction are higher than those of complements attracted to the copulative construction, despite the verbal construction being overall less frequent (see Table 1). This indicates that a relatively small number of complements account for a considerable proportion of all verbal constructions. Indeed, *tov* 'good' alone accounts for 40.4% of verbal constructions with *nire* 'look', and the top

5 complements taken together account for 77.8%. In stark contrast, no single complement of *nire* accounts for even 2% of copulative constructions.

Table 3. Top 5 complements of *nishma* 'sound' in copulative and verbal constructions.

Preference for copulative construction				**Preference for verbal construction**			
Complement	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength	Complement	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength
hegyoni 'logical'	385 (289.88)	2 (97.12)	46.27	*tov* 'good'	57 (576.01)	712 (192.99)	∞
mukar 'familiar'	193 (149.81)	7 (50.19)	16.34	*metsuyan* 'excellent'	18 (168.53)	207 (56.47)	104.40
mufrax 'unfounded'	130 (98.12)	1 (32.88)	14.96	*nehedar* 'terrific'	5 (106.36)	137 (35.64)	76.04
meanyen 'interesting'	114 (86.14)	1 (28.86)	12.97	*axer(et)* 'different'	32 (109.36)	114 (36.64)	41.23
pashut 'simple'	98 (73.41)	0 (24.59)	12.40	*nifla* 'wonderful'	12 (56.18)	63 (18.82)	26.22

Table 4. Top 5 complements of *meriax* 'smell' in copulative and verbal constructions.

Preference for copulative construction				**Preference for verbal construction**			
Complement	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength	Complement	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength
amiti 'real'	1 (0.04)	0 (0.96)	1.40	*tov* 'good'	0 (4.70)	118 (113.30)	2.96
muxan 'ready'	1 (0.04)	0 (0.96)	1.40	*ra* 'bad'	0 (1.19)	30 (28.81)	0.57
mutslax 'successful'	1 (0.04)	0 (0.96)	1.40	*nifla* 'wonderful'	0 (0.96)	24 (23.04)	0.45
mefukpak 'dubious'	1 (0.04)	0 (0.96)	1.40	*nehedar* 'terrific'	0 (0.44)	11 (10.56)	0.20
naki 'clean'	1 (0.04)	0 (0.96)	1.40	*axer(et)* 'different'	0 (0.24)	6 (5.76)	0.11

Turning to complements of the other CPVs, the exact same patterns emerge; see Tables 3-5 for *nishma* 'sound', *meriax* 'smell' and *margish* 'feel', respectively. Complements attracted to the verbal construction are all subjective and, with the exception of *naxon* 'correct', can be predicated of anything. Moreover, these complements are extremely uniform, both within each verb – with very few complements accounting for considerable proportions of all verbal constructions; and across verbs – with the same 8 predicates recurring as the top 5 complements of all 4 CPVs. In comparison, complements attracted to the copulative construction are more varied both within and across verbs, with only 3 predicates recurring among the top 5 complements of the 4 CPVs. Again, many of these complements are not subjective, and many can only be predicated of specific classes of objects, e.g., *mufrax* 'unfounded' for abstract ideas and *naki* 'clean' for physical objects.

Table 5. Top 5 complements of *margish* 'feel' in copulative and verbal constructions.

Preference for copulative construction				**Preference for verbal construction**			
Complement	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength	Complement	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength
amiti 'real'	6 (4.80)	0 (1.20)	0.59	*tov* 'good'	1 (15.21)	18 (3.79)	12.53
zar 'unfamiliar'	5 (4.00)	0 (1.00)	0.49	*naxon* 'correct'	17 (24.82)	14 (6.18)	3.23
gadol 'large'	5 (4.00)	0 (1.00)	0.49	*muzar* 'strange'	2 (6.41)	6 (1.59)	3.00
kaved 'heavy'	5 (4.00)	0 (1.00)	0.49	*axer(et)* 'different'	0 (2.40)	3 (0.60)	2.12
xashuv 'important'	4 (3.20)	0 (1.00)	0.39	*nifla* 'wonderful'	1 (2.40)	2 (0.60)	0.99

The top 5 subjects attracted to each of the copulative and verbal constructions, for each of the four CPVs, are presented in Tables 6-9. As Table 6 below shows, all but one of the subjects of *nire* 'look' attracted to the verbal construction denote objects which can be directly perceived via vision, e.g., *baxur* 'young person'. In contrast, subjects attracted to the copulative construction are just as likely to denote objects which cannot be perceived in the same way, e.g., *she'ela* 'question'. Objects not perceivable through the relevant sensory modality are shaded in gray in Tables 6-9 below. It should be noted that *davar* 'thing', *ish* 'person' and *xayim* 'life' are attracted to the verbal

construction despite having more occurrences in the copulative construction. As high as their observed frequencies in the copulative construction are, they are nevertheless lower than their expected frequencies, and vice versa for the verbal construction.

Table 6. Top 5 subjects of *nire* 'look' in copulative and verbal constructions.

Preference for copulative construction				**Preference for verbal construction**			
Subject	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. Strength	Subject	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength
derex 'way'	153 (126.36)	2 (28.64)	11.02	*davar* 'thing'	679 (912.23)	440 (206.77)	65.19
she'ela 'question'	82 (66.85)	0 (15.15)	7.30	*baxur* 'young person'	20 (53.80)	46 (12.20)	18.98
tshuva 'answer'	62 (50.54)	0 (11.46)	5.51	*kvutsa* 'group'	29 (59.51)	44 (13.49)	14.59
matara 'goal'	49 (39.95)	0 (9.05)	4.36	*ish* 'person'	236 (293.48)	124 (66.52)	12.68
nekuda 'point'	60 (49.73)	1 (11.27)	4.25	*xayim* 'life'	217 (259.24)	101 (58.76)	8.23

Data for the other CPVs again match the pattern observed with *nire* 'look'; see Tables 7-9 below for *nishma* 'sound', *meriax* 'smell' and *margish* 'feel', respectively. In fact, all of the subjects attracted to the verbal construction denote objects directly perceivable through the relevant sensory modality: hearing, smell, or bodily experience. In contrast, several subjects attracted to the copulative construction denote objects which cannot be perceived in the same way, e.g., *tsura* 'shape' via hearing, *sipur* 'story' via smell, and *she'ela* 'question' via touch. It may also be noted that, as Table 7 shows, subjects of *nishma* attracted to the verbal construction tend to denote musical instruments and audio devices, e.g. *(kli) neshifa* 'wind instrument' and *oznit* 'earphone', whereas subjects attracted to the copulative construction tend to denote linguistic objects, e.g., *safa* 'language'.

Table 7. Top 5 subjects of *nishma* ‘sound’ in copulative and verbal constructions.

Preference for copulative construction				**Preference for verbal construction**			
Subject	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength	Subject	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength
davar ‘thing’	203 (181.27)	40 (61.73)	3.53	*(kli) neshifa* ‘wind instrument’	4 (10.44)	10 (3.56)	3.42
she'ela ‘question’	26 (19.39)	0 (6.61)	3.33	*muzika* ‘music’	23 (32.82)	21 (11.18)	2.99
safa ‘language’	12 (8.95)	0 (3.05)	1.53	*oznit* ‘earphone’	0 (3.73)	5 (1.27)	2.98
tshuva ‘answer’	15 (11.94)	1 (4.06)	1.23	*ma'arexet* ‘system’	3 (8.21)	8 (2.79)	2.89
tsura ‘shape’	9 (6.71)	0 (2.29)	1.15	*tof* ‘drum’	6 (11.94)	10 (4.06)	2.74

Table 8. Top 5 subjects of *meriax* ‘smell’ in copulative and verbal constructions.

Preference for copulative construction				**Preference for verbal construction**			
Subject	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength	Subject	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength
sipur ‘story’	1 (0.20)	0 (0.80)	0.70	*ish* ‘person’	0 (0.60)	3 (2.40)	0.30
brauni ‘brownie’	1 (0.20)	0 (0.80)	0.70	*xomer* ‘material’	0 (0.40)	2 (1.60)	0.20
bira ‘beer’	1 (0.20)	0 (0.80)	0.70	*mutsar* ‘product’	0 (0.40)	2 (1.60)	0.20
dmut ‘character’	1 (0.20)	0 (0.80)	0.70	*xeder* ‘room’	0 (0.20)	1 (0.80)	0.10
kokteyl ‘cocktail’	1 (0.20)	0 (0.80)	0.70	*netax* ‘chunk’	0 (0.20)	1 (0.80)	0.10

Table 9. Top 5 subjects of *margish* ‘feel’ in copulative and verbal constructions.

Preference for copulative construction				**Preference for verbal construction**			
Subject	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength	Subject	Copula. obs. (exp.)	Verbal obs. (exp.)	Collostr. strength
xayim ‘life’	3 (2.37)	0 (0.63)	0.31	*xavaya* ‘experience’	0 (0.79)	1 (0.21)	0.68
derex ‘way’	2 (1.58)	0 (0.42)	0.21	*xelbon* ‘protein’	0 (0.79)	1 (0.21)	0.68
she'ela ‘question’	2 (1.58)	0 (0.42)	0.21	*xalal* ‘space’	0 (0.79)	1 (0.21)	0.68
ish ‘person’	2 (1.58)	0 (0.42)	0.21	*sadna* ‘workshop’	0 (0.79)	1 (0.21)	0.68
yad ‘hand’	2 (1.58)	0 (0.42)	0.21	*rexiva* ‘riding’	0 (0.79)	1 (0.21)	0.68

1.4 Discussion

Summarizing the corpus findings detailed above, *nire* ‘look’, *nishma* ‘sound’ and *margish* ‘feel’ all occur in copulative constructions more frequently than in verbal constructions, whereas *meriax* ‘smell’ predominantly occurs in verbal constructions. Additionally, *nire* stands out for frequently occurring in impersonal constructions, which are relatively rare for the other CPVs. Across verbs, copulative and impersonal constructions take dative arguments more frequently than verbal constructions. Complements in verbal constructions are uniform, subjective and can be predicated over anything, whereas complements in copulative constructions are more varied, and more likely to be non-subjective and narrower in terms of what they can be predicated over. Subjects in verbal constructions tend to denote perceivable objects, whereas in copulative constructions they can denote non-perceivable objects.

All in all, the corpus data indicate that the verbal and copulative constructions map to attributary and parenthetical meanings, respectively. Recall that on the attributary meaning, CPVs attribute a property to (a perceptual impression of) an object. This requires an object with a relevant perceptual impression, along with a property applicable to that impression (e.g., Petersen & Gamerschlag 2014). Verbal constructions have parallel requirements, in that their subjects denote

perceivable stimuli, and their complements can be predicated of perceptual impressions (since they can be predicated of anything). Moreover, their complements are predominantly subjective, in accordance with previous observations that attributary uses tend to be value judgments (Rogers 1974, Usoniene 2000, Whitt 2009, 2011, Poortvliet 2018, Viberg 2019). The idea that verbal constructions map to attributary meanings also aligns with previous suggestions that complements in attributary uses behave like, have evolved from, or structurally are, adverbs (Postal 1971, Taniguchi 1997, Gisborne 2010).

In addition, we may consider Gisborne's (2010) diagnostic for identifying the attributary meaning: the impossibility of the continuation *but (it) isn't really*. Hebrew verbal constructions are indeed unacceptable with such a continuation, which can be explained by grammatical constraints on ellipsis. Thus, sentences like (9) are ruled out because the elided material can be neither an adverb – which would be missing a verb, nor an adjective – which would be missing an antecedent. Contrast this with the copulative construction in (10), which allows the continuation because the elided adjective has an antecedent.

(9) *raxel niret muzar, aval be'etsem hi lo ~~muzar~~ / ~~muzara~~
Rachel looks.FSG weirdly, but actually she not weirdly / weird.FSG
'Rachel looks weird, but isn't really.'

(10) raxel niret muzara, aval be'etsem hi lo ~~muzara~~
Rachel looks.FSG weird.FSG, but actually she not weird.FSG
'Rachel looks weird, but isn't really.'

On the parenthetical meaning, CPVs take a proposition as an argument. If we assume that this proposition may be formed by predicating the verb's complement over its subject, then the two are only restricted in that the complement must be applicable to the subject, and, possibly, in that the complement must be inferable from the relevant sensory modality (Petersen & Gamerschlag 2014). Copulative constructions show a similar permissiveness, taking subjects which denote stimuli not perceivable through the relevant sense, or at all, as well as complements which cannot be predicated of physical objects or perceptual impressions thereof. The diversity of complements across CPVs may reflect how different properties are inferable from different senses, and the overall paucity of copulative constructions with *meriax* 'smell' may reflect how relatively few properties are inferable from smell, at least for Hebrew speakers.

At this point, we might consider existing analyses of the two constructions. A typical account of the attributary meaning is Petersen & Gamerschlag's (2014) frame-theoretic analysis (for an analysis along the same lines in model-theoretic semantics, see Muñoz 2019). In their analysis, each CPV specifies an attribute (e.g., SOUND, TASTE) of its subject's denotation, and the CPV's complement assigns a value to that attribute (e.g., SOUND: *high*, TASTE: *bitter*). Thus, a felicitous attributary use must satisfy two constraints: the attribute specified by the verb must be appropriate for the subject's denotation, and the value assigned by the complement must be appropriate for that attribute. The former blocks attributary uses with subjects denoting stimuli which are not perceivable via the relevant perceptual modality (e.g., *#the music looks X*, *#the dish sounds Y*), and the latter blocks attributary uses with complements which cannot be predicated of relevant perceptual impressions (e.g., *#X looks logical*, *#Y sounds blue*).

However, if we apply Petersen & Gamerschlag's (2014) analysis to the Hebrew verbal construction, we over-generate which complements the construction takes. If the construction simply requires a complement which can be predicated of the relevant perceptual impression, why is it that complements are overwhelmingly subjective? Granted, some predicates might be ruled out because they have no adverbial form; this could explain why there are no color terms among the complements of the verbal construction, for instance. But there are predicates in Hebrew which can be predicated of perceptual impressions, and do occur as adverbs, yet never occur in the verbal construction; see (11-13).

(11) a. raxel shara xazak /xalash /gavoha /namux.
Rachel sings.FSG strongly/weakly /high /low
'Rachel sings loudly/softly/with a high/low-pitched voice.'

b. *raxel nishma'at xazak /xalash /gavoha /namux.
Rachel sounds.FSG strongly/weakly /high /low
'Rachel sounds loud/soft/high-pitched/low-pitched.'

(12) a. ha-menora meira xazak /xalash.
the-lamp shines.FSG strongly/weakly
'The lamp shines brightly/softly.'

b. *ha-menora niret xazak /xalash
the-lamp looks.FSG strongly /weakly
'The lamp looks bright/soft.'

(13) a. ha-mexonit nosa'at xalak.
the-car drives.FSG smoothly
'The car handles smoothly.'

b. *ha-mexonit margisha xalak.
the-car feels.FSG smoothly
'The car feels smooth.'

Before we proceed, we should clarify what it means for a predicate to be subjective. This issue has received increasing attention in the past two decades, but remains thorny from both an empirical and a theoretical perspective (e.g., Lasersohn 2005, Kennedy 2013, McNally & Stojanoivc 2017). The most widespread diagnostic for identifying subjective predicates is whether they give rise to faultless disagreement, the intuition that speakers can disagree about certain propositions, e.g., whether something is tasty, without any one of them being objectively wrong (Kölbel 2004). However, experimental evidence shows that faultless disagreement arises to different extents for different predicates (Solt 2018) and depends on social consensus as much as on predicate choice (Kaiser & Rudin 2020). Moreover, faultless disagreement arises from at least three different sources: dimensional vagueness, multidimensionality, and judge-dependence (Kennedy 2013, McNally & Stojanovic 2017, Solt 2018).

Dimensional predicates such as *strong* and *high* may allow faultless disagreement due to vague standards. For instance, speakers might disagree about whether or not *Rachel is strong*, due to having different standards for what counts as *strong* in context. Notably, faultless disagreement disappears when these predicates occur in comparative and superlative constructions, e.g., *Rachel is stronger than Adam.*

Multidimensional predicates such as *smart* and *healthy* allow faultless disagreement by involving multiple criteria which can be used to determine whether, and to what degree, an object has the property in question. For instance, speakers might disagree about whether *Rachel is healthier than Adam*, due to having different considerations in determining how *healthy* someone is, e.g., their

blood pressure, immune system, etc. Multidimensional predicates can be identified with phrases which explicitly target their dimensions, e.g., *healthy with respect to X* (Sassoon 2013).

Judge-dependent predicates such as *fun* and *tasty* give rise to faultless disagreement by involving a judgment regarding taste, value, etc. For instance, speakers might disagree about whether *Rachel is more fun than Adam*, because the degree to which something is *fun* is a matter of personal judgment rather than fact. Judge-dependent predicates can often – though not always – be identified by taking an optional argument representing the judge, e.g., *fun for X* (Lasersohn 2005, McNally & Stojanovic 2017).

For present purposes, it is clear that dimensional vagueness alone does not license complements in the verbal construction. As the examples in (11a-13a) above show, the dimensional complements *xazak* 'strong', *xalash* 'weak', *gavoha* 'high', *namux* 'low', and *xalak* 'smooth', are not acceptable in the verbal construction, even in non-comparative forms. As for multidimensionality and judge-dependence, most of the predicates attracted to the complement slot of the verbal construction are both multidimensional and judge-dependent (*tov* 'good', *nehedar* 'terrific', etc.). But are either of these a necessary or sufficient condition for complements in the verbal construction? Paradigmatic multidimensional predicates which are not judge-dependent, *xaxam* 'smart' and *bari* 'healthy', and likewise paradigmatic judge-dependent predicates which are not multidimensional, *kef* 'fun' and *taim* 'tasty', are ruled out for independent reasons – they have no adverbial form, or cannot be predicated of relevant perceptual impressions. There are, however, multidimensional predicates which are not judge-dependent, and which are attracted to the verbal construction, most prominently *axeret* 'different' but also other (dis)similarity predicates such as *dome* 'similar' and *shone* 'different'. In contrast, only one judge-dependent predicate which is not multidimensional – *naxon* 'correct' – is attracted to the verbal construction to a comparable extent, and only with the verb *margish* 'feel'.[16]

[16] My intuition is that *naxon* 'correct' in these cases is used figuratively, to express something like 'appropriate' or 'satisfactory', similar to English *feel right*; see (i):

(i) yesh batim ha-mesudarim be-ofen she-hem "margishim naxon" [g]
there.are houses that-arranged.PL in-way that-they feel.PL correctly
'Some houses are arranged in a way that "feels right".'

It appears, then, that the verbal construction attracts multidimensional predicates to its complement slot. To account for this, I propose that the CPV in the verbal construction interacts directly with the dimensions of its complement. That is, rather than specifying an attribute of its subject's denotation, as suggested by Petersen & Gamerschlag (2014), the CPV restricts the interpretation of its complement to a subset of dimensions: visual dimensions for *nire* 'look', auditory dimensions for *nishma* 'sound', etc. This functions similarly to phrases which directly target a multidimensional predicate's dimensions, e.g., *with respect to X* (Sassoon 2013, see also Alrenga 2010), such that (14a) is analyzed as equivalent to (14b). If the complement is not a multidimensional predicate, the CPV has no dimensions to target and composition fails; see (15a-b). Likewise, if the resulting predicate – restricted to a specific subset of dimensions – cannot apply to the subject's denotation, composition fails; see (16a-b).

(14) a. raxel niret muzar.
Rachel looks.FSG weirdly
'Rachel looks weird.'

b. raxel muzara mi-bxinat mare
Rachel weird.FSG from-aspect.of look
'Rachel is weird with respect to looks.'

(15) a. *ha-menora niret xazak.
the-lamp looks.FSG strongly
'The lamp looks bright.'

b. ha-menora xazaka (*mi-bxinat mare)
the-lamp strong.FSG (from-aspect.of look)
'The lamp is bright (with respect to looks).'

(16) a. *ha-muzika niret muzar.
the-music looks.FSG weirdly
'The music looks weird.'

b. ha-muzika muzara (*mi-bxinat mare)
the-music weird.FSG (from-aspect.of look)
'The music is weird with respect to looks.'

We can see that the CPV restricts its complement to a subset of dimensions, and not to a single dimension, from the possibility of it co-occurring with another phrase targeting a predicate's dimensions. Thus in (17a), the predicate *niret muzar* 'look weird' is restricted to visual dimensions but remains multidimensional, allowing the phrase *mi-bxinat tseva* 'with respect to color' to further restrict its interpretation to a single visual dimension, color. In contrast, (17b) with the phrase *mi-bxinat ta'am* 'with respect to flavor' is unacceptable, because the subset of dimensions which *niret muzar* 'look weird' is restricted to, does not include the dimension of flavor.

(17) a. ha-uga niret muzar mi-bxinat tseva.
the-cake looks.FSG weirdly from-aspect.of color
'The cake looks weird with respect to color.'

b. *ha-uga niret muzar mi-bxinat ta'am.
the-cake looks.FSG weirdly from-aspect.of flavor
'The cake looks weird with respect to flavor.'

The present proposal explains why CPVs in the verbal construction so rarely co-occur with a dative argument in the corpus (between 2.5% and 4.1%). Unlike the copulative and impersonal constructions (see below), the verbal construction does not actually have a semantic role for the dative argument to fill (see Gisborne 2010). In the few cases where the CPV does take a dative argument, it likely fills the judge role required by a judge-dependent complement, e.g., *tov* 'good', *ra* 'bad', which licenses a dative argument on its own.

Moving on to the copulative construction, the CPV appears to function as a raising verb: its subject and complement combine to form a proposition, which the CPV takes as an argument and modifies (see also Avineri 2021, who analyzes the complement in this construction as a small clause). The present finding that the copulative construction attracts abstract subjects lends credence to analyses which don't require the subject to be a source of evidence (Gisborne 2010, cf. Muñoz 2019). The remaining question is whether CPVs in the copulative and impersonal constructions are semantically equivalent. Particularly, does the CPV in each construction have an evidential meaning, an epistemic modal meaning, or both?

Note that I use the term evidentiality to refer to any encoding of information about the source or type of evidence for a proposition, whether or not it is grammaticized. I distinguish a priori between

evidentiality and epistemic modality, which refers to the encoding of an individual's beliefs or knowledge. It's possible, however, for an expression to encode both evidential and epistemic meaning, and in fact, Matthewson (2012) explicitly claims that every expression which encodes one also encodes the other. Even if an expression only encodes evidential meaning, its use may convey epistemic information via pragmatic inference, and vice versa (Degen et al. 2019). This makes it difficult to tease apart evidential and epistemic meaning within corpus data. Thus, to test whether the copulative and impersonal constructions differ in encoding evidential and epistemic meaning, I conducted a preference experiment.

1.5 Experimental study

The experiment was designed to test whether copulative and impersonal constructions differ in encoding evidential and epistemic meanings. I restricted the experiment to *nire* 'look' and *nishma* 'sound', since *meriax* 'smell' and *margish* 'feel' occurred in impersonal constructions very rarely in the corpus. The experiment relied on contradictions to the evidential and epistemic meanings presumably encoded by the CPV, specifically direct perception of an object (e.g., Landau 2011) and compatibility with the speaker's knowledge (e.g., Hansen & Markman 2005). I assumed that, if one construction encoded evidential meaning but the other did not, then the latter would be preferred in contexts contradicting the evidential meaning. Likewise, if one construction encoded epistemic meaning but the other did not, then the latter would be preferred in contexts contradicting the epistemic meaning.

The experiment consisted of categorical preference judgments. In each trial, participants were presented with a short context introducing two speakers discussing an object, and an incomplete sentence by one of the speakers. Participants were asked to indicate which of two possible continuations would be better, that is, would complete the speaker's sentence in a way that sounds more natural, logical and reasonable. Participants could also choose to indicate that neither continuation was better than the other.

In every trial, the presented incomplete sentence contradicted either the evidential or the epistemic meaning presumably encoded by the CPV. For the evidential meaning, the incomplete sentence asserted that the speaker did not directly perceive the object under discussion via the sensory modality associated with the CPV (seeing for *nire* 'look', hearing for *nishma* 'sound'). For the epistemic meaning, the incomplete sentence asserted that the speaker had prior knowledge

contradicting the proposition argument of the CPV. In the critical trials, the two continuations presented to participants were one copulative and one impersonal construction, both with the same CPV. To ensure that a contradiction arises, all continuations included a 1st person dative argument. Examples of critical trials, translated from Hebrew, are given in (18-19) (in the original Hebrew, the dative argument immediately follows the CPV in both constructions, and there is no expletive subject in the impersonal construction). The full list of materials is available online at https://osf.io/g7kw6/.

(18) Noa is consulting with Hila about Professor Sabag's course.

Hila: "I've never heard her, but…	(No direct perception)
… she sounds nice to me."	(Copulative)
… it sounds to me that she's nice."	(Impersonal)

(19) Dvir and Ehud are gossiping about the lecturer in Statistics for Historians.

Dvir: "I know that she's young, but…	(Prior knowledge)
… she sounds old to me."	(Copulative)
… it sounds to me that she's old."	(Impersonal)

A total of 32 critical trials were created, half with no direct perception, and half with prior knowledge. Half of each included *nire* 'look' in their continuations, and the other half included *nishma* 'sound'. 16 filler trials were also created, in which the two continuations presented to participants were instances of the same construction, with two different CPVs. The experiment was created in Google Forms, and participants were recruited via social media. 41 native speakers of Hebrew volunteered to participate. Before completing the experiment, participants saw two example trials which did not include CPVs, and were advised that there were no right or wrong answers. Each participant then saw all 48 trials, in a pseudo-random order (the same condition did not appear twice in a row).

The results of the preference experiment are presented in Figure 2. In trials where the speaker had prior knowledge contradicting the proposition argument of the CPV, participants consistently preferred the copulative construction, both with *nire* 'look' (76.8%) and with *nishma* 'sound' (80.5%). In trials where the speaker did not directly perceive the object in question, speaker preferences depended on the CPV in the available continuations. With *nire*, speakers generally

preferred the impersonal construction (67.8%), but with *nishma*, there was no preference for either construction (35.1% copulative, 30.5% impersonal, 34.5% no preference).

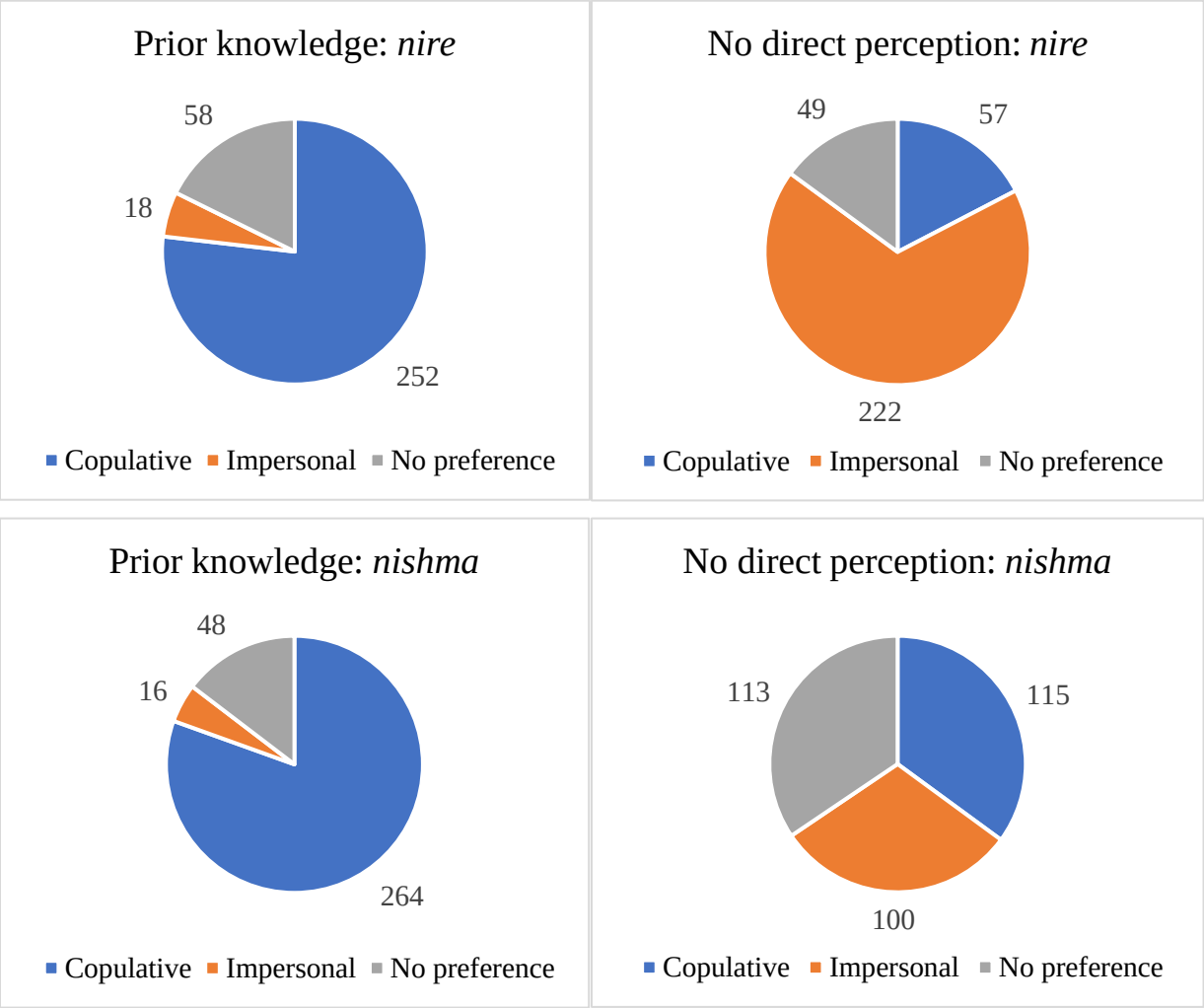


Figure 2. Results of the preference experiment. Participants chose whether the copulative or impersonal construction was better in trials where the speaker had prior knowledge contradicting the proposition argument of the CPV, and in trials where the speaker did not directly perceive the object in question.

These results strongly suggest that CPVs encode compatibility with the speaker's beliefs – an epistemic modal meaning – in the impersonal construction, but not in the copulative construction. The impersonal construction is consistently dispreferred when the speaker's knowledge contradicts the CPV's proposition argument, similarly to how epistemic *might* or *must* are unacceptable when their prejacent contradicts the speaker's knowledge.

In contrast, the results suggest that *nire* 'look' encodes direct perception of the subject – evidential meaning – in the copulative construction, but not in the impersonal construction. The copulative construction is consistently dispreferred when the speaker did not directly perceive the object in question, regardless of whether they believe what their perception evidences. Note that direct perception of the subject is a default assumption rather than a logical entailment, given the corpus finding that the copulative construction attracts abstract subjects which cannot be seen (e.g., *she'ela* 'question', *matara* 'goal'). Additionally, the results do not preclude the impersonal construction from encoding a weaker evidential meaning, e.g., visual evidence from an unspecified source.

The picture is not as clear-cut for *nishma* 'sound'. Since there was no preference for either construction in trials where the speaker did not hear the object in question, it's possible that both constructions were acceptable, or that both constructions were unacceptable. I hypothesize that participants found both constructions acceptable, not because *nishma* lacks evidential meaning, but because it allows indirect reported evidence in addition to direct auditory perception (see also Gisborne 2010, Landau 2011, Viberg 2019).

These results raise a possible explanation for why the impersonal construction occurs in the corpus almost exclusively with *nire* 'look'. If the meaning of the impersonal construction is primarily epistemic rather than evidential, it essentially doesn't matter which CPV is used. Thus *nire*, as the most frequent CPV to begin with, might be used by default.

1.6 Conclusions

In this chapter I investigated Hebrew CPVs occurring in the verbal, copulative and impersonal constructions. The alternation between the verbal and copulative constructions has not been discussed in previous literature, although it appears in at least one language other than Hebrew (see fn. 5). I presented corpus data suggesting that the formal distinction between the verbal and copulative constructions corresponds to the semantic distinction between attributary and parenthetical meanings of CPVs.

The corpus data additionally show that CPVs in the verbal construction attract subjective multidimensional predicates. This weakens existing accounts of the attributary meaning, which predict that it may take any complement that can be predicated over the relevant type of perceptual impression (Gisborne 2010, Petersen & Gamerschlag 2014, Muñoz 2019). I therefore proposed that CPVs in the verbal construction should be given a different analysis, one which interacts

directly with the dimensions of the complement, along the lines of Alrenga (2010) and Sassoon (2013). Note that this proposal is compatible with the observation that CPVs may also have an attributary meaning with comparative complements, since comparative phrases too are multidimensional (Alrenga 2010).

I would go further and suggest that in languages other than Hebrew, attributary uses may be more restricted than previously assumed, in ways similar to the Hebrew verbal construction. Specifically, I believe we should reexamine cases where CPVs take complements associated with a single sensory modality, e.g. *look oblong*, *sound loud*, *feel sticky*. Such cases are ascribed attributary meanings by Gisborne (2010) and Petersen & Gamerschlag (2014), and yet Gisborne's (2010) own diagnostic for identifying the attributary meaning – the impossibility of the continuation *but (it) isn't really* – suggests that they can in fact have a parenthetical meaning; see (20a-c).

(20) a. It **looks oblong, but** we know the moon **is really round**. [g]
b. a *really* good drummer can balance their playing so that it **sounds loud but isn't**. [g]
c. Playfoam is a great sensory moulding material which **feels sticky but isn't**! [g]

Moreover, there is no apparent communicative function for an attributary use of a CPV taking a complement associated with a single sense. To borrow Petersen & Gamerschlag's (2014) terminology, if the phrase *sound loud* simply assigned a value to the SOUND attribute of the subject's denotation, then it would be equivalent to the bare predicate *loud*, given that SOUND is the only attribute that *loud* can be assigned to in the first place. In other words, a CPV only makes a semantic contribution with a predicate that can in principle assign values to more than one attribute, i.e., a multidimensional predicate. It remains to be seen whether attributary uses in other languages actually follow this logic, but if they do, we would expect them to predominantly take multidimensional complements, just like CPVs in the Hebrew verbal construction.[17]

I also conducted a preference experiment comparing the semantics of the copulative and impersonal constructions. The results indicate that the copulative construction encodes a more

[17] If attributary uses in other languages turn out to be restricted to multidimensional complements, it would also explain the unacceptability of "goal-oriented" complements, e.g., *#This dish tastes good (for making people queasy)*, which should be possible if the predicate applies, unmodified, to a perceptual impression or attribute of the subject (see Muñoz 2019).

specific evidential meaning than the impersonal construction, namely direct perception of the CPV's subject. The impersonal construction, conversely, encodes an epistemic modal meaning, namely compatibility of the proposition argument with the speaker's knowledge. The former finding aligns with previous claims that the impersonal construction doesn't specify a source of evidence (e.g., Landau 2011, Muñoz 2019), and the latter with Hansen & Markman's (2005) suggestion that the impersonal construction can only refer to "likely reality" and not to "outward appearance". More generally, these findings reinforce the need to maintain a distinction between the class of evidentials and the class of epistemic modals (cf. Matthewson 2012).

Taken together, the empirical findings in this chapter highlight the semantic and syntactic versatility of CPVs. This versatility opens up opportunities for future research, but also presents potential pitfalls. By virtue of their threefold semantics – encompassing the domains of perception, subjectivity, evidence, and belief – CPVs can be used to investigate a wide range of topics in linguistics and neighboring disciplines. However, any work that uses CPVs in this way must acknowledge their multiple meanings and take steps to keep those meanings distinct, which is not always trivial.[18] In the remainder of this dissertation I endeavor to do just that, for two separate case studies, implementing the lessons of the present chapter.

In contrast to the present chapter's focus on Hebrew, the next two chapters are based entirely on English data. The reasons for this shift are predominantly matters of practicality. Chapter 2 relies on modality strength norms elicited for English adjectives (Lynott & Connell 2009), which have no Hebrew counterpart. Chapter 3 explores the idea of a hierarchy of the 5 classical senses, and Hebrew has no established CPV for the sensory modality of taste (see fn. 11). Additionally, chapter 3 includes a corpus study exploring a very rare construction, occurrences of which are even harder to find in Hebrew than in English (I was able to find fewer than 10 relevant occurrences on the Hebrew Web 2014 corpus). Both chapters also include experimental studies, and English-speaking participants are substantially easier to recruit than Hebrew-speaking participants. Finally, for better

[18] This can be especially tricky if the research topic naturally calls for CPVs with multidimensional predicates as complements, since – in light of present findings – such instances are predicted to allow both attributary and parenthetical meanings. This would be a challenge in amending McNally & Stojanovic's (2017) diagnostic for identifying aesthetic and evaluative adjectives, many of which are multidimensional. Similarly, it would be a challenge in replicating Kaiser's (2018) experiment on the effect of sensory modality on the identification of a subjective attitude holder, as the experimental design requires multidimensional predicates if the materials are to be uniform across sensory modalities.

or worse, English data remain more accessible and likely of greater interest to the academic community at large. Nevertheless, the findings of the present chapter will prove relevant in the investigations undertaken in the following chapters. I hope other researchers might find them useful as well.

2 Evidential uncertainty

2.1 Introduction

One of the principal uses of language is to exchange information about the state of the world, but this information is not always free from doubt or reservation. Speakers might only have indirect or unreliable evidence for what they believe, and a cooperative speaker may seek to convey this to their audience. Natural languages offer various means of making one's evidence explicit, and interlocutors can reason about each other's evidence whenever it is left implicit. This chapter is part of a growing effort by semanticists and pragmaticists to tackle these phenomena, theoretically and empirically (e.g. Lassiter 2016, Pogue & Tanenhaus 2018, Degen et al. 2019, Ünal & Papafragou 2020). It focuses on how the use of evidential expressions interacts with inferences about belief and perception.

In conversation, we routinely assume that other speakers believe what they say, and moreover, that they have adequate evidence to support their beliefs (Grice 1989). In about a quarter of the world's languages, it is in fact obligatory to mark what type of evidence a speaker has for an assertion, just as it is obligatory to mark tense on an English verb, or gender on a Hebrew adjective (Aikhenvald 2004, 2014). But most languages do not have obligatory grammatical marking of evidence. Thus, an English speaker who utters (21) below is assumed to believe that the dress is new, and to have evidence in support of that belief, even though the nature of this evidence is left implicit.

(21) The dress is new.

Speakers in any language may, however, choose to explicitly indicate what type of evidence they have, using evidential expressions such as perception verbs, parentheticals and modals. To illustrate, (22a) indicates by default that the dress' appearance evidences that it is new (see chapter 1), (22b) indicates reported evidence (Simons 2007), and (22c) indicates an indirect inference (von Fintel & Gillies 2010, Mandelkern 2019). The use of such evidential expressions carries an additional discourse effect, namely conveying that the speaker is not entirely confident in what they assert. For instance, the speakers of (22a-c) would be perceived as less confident that the dress is new, relative to the speaker of (21) (see experimental evidence in Pogue & Tanenhaus 2018, Degen et al. 2019).

(22) a. The dress looks new.

b. The dress is new, I hear.

c. The dress must be new.

Why is it that explicitly marking what type of evidence one has conveys lower confidence in one's beliefs, compared to leaving that evidence implicit? This is the main question that the current chapter aims to answer. I identify two approaches to explaining the reduced confidence associated with evidential expressions. The first approach, extrapolated from von Fintel & Gillies' (2010) work on *must* (see also Mandelkern 2019), attributes this effect to extralinguistic reasoning about the type of evidence indicated by the speaker, specifically about its (in)directness. To illustrate, in each of the utterances in (22a-c), the indicated type of evidence only provides indirect support for the belief that the dress is new. Hence, none of that evidence justifies high confidence in that proposition.

A second approach, put forth by Degen et al. (2015), attributes the reduced confidence to Gricean reasoning about alternative utterances the speaker could have used, but didn't (Grice 1989). For example, the speaker of any of the utterances in (22a-c) could have used the shorter, and thus less costly, bare utterance in (21) instead. That would have conveyed high confidence that the dress is new. Listeners then reason that the speaker must have had a reason to go to the effort of using the longer and costlier utterance, and one such reason could be to avoid conveying the same high degree of confidence.

These two approaches make conflicting predictions regarding a particular set of utterances so far largely overlooked by researchers, namely ones in which an evidential expression marks the most direct type of evidence possible for a proposition. Consider (23a) below, in which the speaker uses the verb *looks* to indicate that they have visual evidence that the dress is blue. On the face of it, this is the most direct type of evidence possible to support a belief about color, a visual property. Therefore, if listeners infer the speaker's degree of confidence from the directness of their evidence – using extralinguistic reasoning – then (23a) should convey high confidence, comparable to (23b). In contrast, if listeners compare (23a) to the less costly alternative in (23b) then, via Gricean reasoning, (23a) should convey reduced confidence, corresponding to the effect observed for (22a-c).

(23) a. The dress looks blue.

b. The dress is blue.

Further complicating the picture is an observation by Grice (Grice & White 1961), dubbed the 'doubt-or-denial' condition, regarding utterances like (23a). Purportedly, (23a) is only felicitous in a context in which the proposition 'The dress is blue' is in doubt or has been denied. Thus, utterances marking direct evidence seem to carry yet another discourse effect, beyond the "ordinary" reduced confidence associated with utterances marking indirect evidence.

I propose a hybrid account of the discourse effects of evidential expressions, which relies on both Gricean reasoning and extralinguistic reasoning about type of evidence. On my proposal, all utterances marking any type of evidence are compared with alternatives without such marking. This routinely triggers an inference of reduced confidence via Gricean reasoning, regardless of the type of explicitly marked evidence. For indirect evidence, this reduced confidence aligns with listeners' extralinguistic assumptions, since world knowledge dictates that indirect evidence coincides with low speaker confidence. But for direct evidence, an additional step of extralinguistic reasoning is required, because direct evidence is normally expected to coincide with high confidence. To solve this clash in expectations, listeners reason that the direct evidence in these particular circumstances must be somehow compromised, thereby deriving the doubt-or-denial condition. My proposal therefore makes two concrete predictions: (i) utterances explicitly marking direct evidence convey reduced confidence relative to bare utterances, and (ii) direct evidence is more likely to be explicitly marked under circumstances where that particular type of evidence is compromised, and hence less reliable, e.g., poor lighting in the case of visual evidence.

The rest of this chapter is structured as follows. In section 2.2, I outline the three aforementioned approaches to the reduced confidence associated with evidential expressions, and spell out their predictions regarding utterances marking direct evidence. In section 2.3, I present a set of experiments empirically testing these predictions, based on the experimental paradigm developed by Degen et al. (2019). In section 2.4, I develop a computational model of the reasoning processes involved in using and interpreting the relevant utterances. My model extends the Rational Speech Act framework (Frank & Goodman 2012, Goodman & Frank 2016) with formal representations of evidence directness and evidence reliability. I conclude with section 2.5.

2.2 Background

This section outlines three accounts for the reduced confidence conveyed by non-obligatory evidential expressions, here called (a) the extralinguistic account, (b) the Gricean account, and (c) the hybrid account. I spell out the predictions of each account regarding utterances marking direct perceptual evidence, e.g., *The dress looks blue*.

The extralinguistic account is a straightforward extrapolation of von Fintel & Gillies' (2010) work on epistemic *must* and a puzzle associated with it, which they dub Karttunen's Problem. This refers to Karttunen's (1972) observation that uttering (24a) conveys that the proposition denoted by in (24b) is not yet an established fact, and that the evidence for it is somehow fallible. This intuition has been confirmed in later experimental work, showing that uttering *must p* conveys a weaker belief in the truth of *p* relative to uttering the bare prejacent *p* (Degen et al. 2015, 2019). The puzzle is how to reconcile these data with the traditional semantics of epistemic *must*, which dictates that *must p* is no less logically strong than *p*.

(24) a. John must have left.

b. John has left.

The standard approach to Karttunen's Problem is to revise the traditional semantics of epistemic *must*, in such a way as to make *must p* logically weaker than *p* (e.g., Kratzer 1991). The weaker belief conveyed by *must p* can then be derived straightforwardly as a scalar implicature: the speaker could have uttered *p* but chose to utter the weaker alternative *must p*, signaling that they don't know that *p* is true. However, von Fintel & Gillies argue against this approach, insisting that epistemic *must* ought to remain logically strong. Instead, their solution to Karttunen's Problem involves adding an evidential component to the meaning of epistemic *must*, motivated by the contrast between (25b) and (26b) (von Fintel & Gillies 2010, pp. 353-4).

(25) [Context: seeing the pouring rain]

a. It's raining.

b. ??It must be raining.

(26) [Context: seeing wet rain gear and knowing rain is the only possible cause]

a. It's raining.

b. It must be raining.

The contrast between (25b) and (26b) demonstrates that epistemic *must* is infelicitous when the speaker's evidence for the prejacent is direct observation rather than indirect inference. To capture this, von Fintel & Gillies propose that epistemic *must* (across languages) presupposes that the speaker does not have information that directly settles the prejacent.[19] Taken together with the traditional semantics of epistemic *must,* the result is that the speaker of (26b) is just as strongly committed to the prejacent as the speaker of (26a), but how committed they are is divorced from how "convincing" they are, to other interlocutors. As von Fintel & Gillies put it, "there are prejacents for which intuitively direct evidence is more convincing evidence… we may judge that in many cases, *must* ϕ is more likely to be false than ϕ by itself would have been if there had been direct evidence for the prejacent" (p. 380).[20]

Under this account, Karttunen's Problem is effectively thrust out of the purview of semantics and pragmatics, and into the realm of extralinguistic reasoning about evidence strength, i.e., about how likely different states of the world are, given various types of evidence. The same theoretical move can be used to account for the reduced confidence conveyed by any evidential expression that marks indirect evidence for a proposition. Such an extralinguistic account, taken broadly, would predict that confidence judgments are based not (only) on the speaker's commitment, i.e., on the logical strength of the utterance, but (also) on the strength of the speaker's evidence.

To probe this broad prediction, we would need test cases with a mismatch between logical strength and evidence strength: either high logical strength with low evidence strength, or low logical strength with high evidence strength. One potential test case is occurrences of *see* with a small

[19] Von Fintel & Gillies propose two separate formalizations of this idea. On the first, the speaker has information that directly settles the prejacent if the speaker knows a single proposition that either entails or contradicts the prejacent. So *must p* can be felicitously and truthfully uttered only if the speaker knows a set of propositions K such that no single proposition $q \in K$ entails p, but the intersection $\cap K$ does entail p. On the second formalization, the context determines a partition over worlds, representing a contextually relevant "set of issues". The speaker has information that directly settles the prejacent if all the worlds in each cell of the partition agree on the prejacent. So *must p* can be felicitously uttered only if *p* isn't one of these contextually relevant issues.

[20] In the elided part of this quote, von Fintel & Gillies write that "a speaker who chooses nevertheless to use the strong *must* ϕ incurs a higher degree of risk", referring to the risk of failing to convince one's interlocutors of a proposition. But of course, the opposite is also a risk: convincing your audience of a proposition that later turns out to be false risks loss of face and trust.

clause complement, which have been argued to logically entail their proposition complement (e.g., Higginbotham 1983); see (27a). To pair this high logical strength with low evidence strength, we would need a small clause referring to an event/situation that is difficult or impossible to see, as in (27b). The extralinguistic account predicts (27b) would convey lower confidence in its proposition complement than (27a), because the speaker of (27b) is more likely to be wrong about what they saw than the speaker of (27a).

(27) a. I saw Mary solve the equation on the whiteboard. (⇒ Mary solved the equation on the whiteboard.)

b. I saw Mary solve the equation in her head. (⇒ Mary solved the equation in her head.)

A second test case is evidential uses of copulative perception verbs, which are logically independent from the proposition they provide evidence for, as shown in (28a-b) (see also chapter 1). To pair this logical weakness with high evidence strength, we would need a proposition whose truth value can be directly perceived via the relevant sensory modality (in von Fintel & Gillies' (2010) terms, a proposition that is directly settled by the relevant sensory perception), as in (29a-c). An extralinguistic account predicts that (29a-c) would convey higher confidence than (28a), because the speakers of (29a-c) are more likely to be right about what they perceived than the speaker of (28a). Experiment 1, described in section 2.3, is designed to test this prediction.

(28) a. The dress looks new (but it isn't).

b. The dress doesn't look new (but it is).

(29) a. The dress looks blue.

b. The dress feels soft.

c. The music sounds loud.

An alternative to the extralinguistic account is the Gricean account put forth by Degen et al. (2015). They sought to capture, in a uniform way, how *must p*, *might p*, and *probably p* each convey a weaker belief in *p* than the bare utterance *p*. Their account is Gricean, in that it relies on pragmatic reasoning about alternative utterances a speaker could have used, but chose not to (Grice 1989). The explanation proceeds as follows: any utterance with an (optional) epistemic modal, evidential, or hedging expression is more costly to the speaker than a bare utterance without such an

expression.[21] Hence, using any utterance with such an expression flaunts the maxim of manner, which requires that speakers "be brief" (Grice 1989). Interlocutors reason that a rational speaker would have a specific intention in flaunting the maxim, namely seeking to communicate information other than the at-issue proposition. Among other things, this information could be that the speaker holds a weaker than default commitment to the truth of the at-issue proposition.

Degen et al.'s Gricean account is fleshed out computationally in the Rational Speech Act framework (Frank & Goodman 2012; introduced in detail in section 2.4 below). While the computational model's precise predictions depend on a number of background assumptions, it can in principle correctly derive that *must p* conveys weaker belief in *p* than the bare utterance *p*.[22] Importantly, it can derive this inference without assuming that *p* is logically stronger than *must p*, unlike previous analyses which derive it as a scalar implicature (e.g., Kratzer 1991). Therefore, Degen et al.'s account can extend to a wide range of evidential expressions, including copulative perception verbs.

As implemented by Degen et al. (2015), the Gricean account includes a limited representation of extralinguistic reasoning, specifically about the strength of evidence available to the speaker, which is related via conditional probability to the speaker's beliefs. Thus, listeners infer that a speaker uttering *must p* not only holds a weaker belief in *p* than a speaker uttering bare *p*, but also that the former has weaker evidence for their belief than the latter. However, the Gricean account does not represent different types of evidence, and therefore cannot capture differences in evidence strength dependent on evidence type. To illustrate, the utterances in (29a-c) above are intuitively alike, in that each of them evokes the strongest possible type of evidence for a proposition. They intuitively differ from the utterances in (30a-c), which mix and match those same evidence types and propositions. The Gricean account cannot capture these intuitions, and consequently does not predict a systematic difference between these sets of utterances.

[21] Costliness is informally described as a function of markedness in Degen et al. (2015) and of length in Degen et al. (2019), but frequency, salience, and complexity all potentially factor into it as well. However exactly it is computed, I find the assumption that bare utterances are less costly than the relevant alternatives to be eminently reasonable.

[22] As implemented in Degen et al. (2015), the model uses unordered threshold semantics for *must p* and bare *p*, inspired by Lassiter & Goodman (2013). Mathematical representations of evidence strengths, speaker beliefs, and utterance costs are either drawn directly from or optimized relative to experimental data collected for the same study.

(30) a. The dress looks soft.

b. The music feels loud.

c. The dress sounds blue.

The extralinguistic account and the Gricean account, as described above, make contrasting predictions regarding certain sets of utterances, e.g., (29a-c) versus (28a) and (30a-c). But the two accounts are not actually mutually exclusive. I propose to integrate them into a hybrid account which involves both extralinguistic reasoning about different types of evidence, and Gricean reasoning about alternative utterances and their costliness. There are potentially multiple different ways to implement such a hybrid account. My own proposal is inspired by Grice's observation that utterances like (31a) below are only felicitous in a context where the proposition they provide evidence for, here denoted by (31b), is in doubt or has been denied (Grice & White 1961). Grice dubs this the doubt-or-denial condition. Supposedly, utterances like (32a) are not subject to the same condition.

(31) a. The dress looks blue.

b. The dress is blue.

(32) a. The dress looks new.

b. The dress is new.

If Grice's observation is right, then the speaker of (31a) may be judged to hold a weaker belief in the evidenced proposition than the speaker of (32a). This is not a result predicted by either of the two accounts outlined above. The extralinguistic account predicts the exact opposite, that the speaker of (31a) would be judged to hold a stronger belief in the evidenced proposition, because they have more direct evidence and are therefore less likely to be wrong. The Gricean account does not predict a systematic difference between the two utterances, because it does not distinguish between types of evidence.

For the hybrid account to predict that (31a) conveys a weaker belief in the evidenced proposition than (32a), two assumptions must hold. First, evidence type must be taken into consideration in determining evidence strength, and consequently in inferring the strength of a speaker's belief. Paradoxically, however, the inferred strength of the speaker's belief must not be directly

proportional to evidence strength, as determined by evidence type; there must be an inflection point, beyond which an ostensible increase in evidence strength actually decreases inferred speaker belief. To reconcile these two conflicting assumptions, I propose to split evidence strength into two components: directness and reliability.

Evidence directness is a relation between types of evidence and propositions, and could be seen as a gradable generalization of von Fintel & Gillies' (2010) categorical directness. Directness represents the intuitive difference between utterances like (29a-c) and (31a) on the one hand, and ones like (30a-c) and (32a) on the other. The evidence types marked by the former utterances attest more directly to the evidenced propositions (e.g., visual evidence to the proposition 'The dress is blue') than the ones marked by the latter (e.g., visual evidence to the proposition 'The dress is new').

Evidence reliability is a relation between types of evidence and contexts. It represents whatever circumstantial factors might cause someone to doubt or deny a proposition even while holding evidence of a relevant type. For example, ambiguous illumination may cause two people directly observing the same picture of a dress, from the same distance at the same angle, to disagree as to whether the dress is blue or yellow (Gegenfurter et al. 2015, Winkler et al. 2015).

With this more nuanced conception of evidence strength, we can explain how marking highly direct evidence might actually decrease inferred speaker belief. The mere use of an evidential expression, rather than a bare utterance, leads listeners to infer that the speaker's evidence is not maximally strong. Since the indicated evidence is highly direct, listeners reason that its weakness must be due to unreliability, i.e., the existence of some contextual reason to doubt or deny the evidenced proposition. Finally, depending on the precise mathematical relation between evidence strength, directness, and reliability, listeners could end up gauging the speaker's evidence strength as weaker than it would be with evidence of moderate directness but high reliability.

2.3 Experimental studies

I present two experimental studies based on the paradigm developed by Degen et al. (2019) for the investigation of evidential expressions. Experiment 1 measures the inferred beliefs of a speaker using bare forms and using evidential expressions implying evidence of varying directness. Experiment 2 explores the choice between using bare forms and using evidential expressions

implying evidence of varying directness, in contexts with either good or poor perceptibility conditions.

In both experiments, evidence directness is based on sensory modality strength norms for adjectives collected by Lynott and Connell (2009). Lynott and Connell generated modality strength norms for the five classical senses for 423 adjectives, by asking participants how strongly a property was experienced by seeing, hearing, feeling through touch, etc. For any given adjective, I take its sensory modality strength as equivalent to the directness of evidence obtained through the relevant sense for a proposition about that adjective. To illustrate, *blue* has a higher visual strength than *new*, hence visual evidence is more direct for the proposition *the dress is blue* than for the proposition *the dress is new*. Consequently, *the dress looks blue* implies more direct evidence than *the dress looks new*.

2.3.1 Experiment 1: Interpretation

This experiment investigates how the directness of evidence available to the speaker, as implied by use of an evidential expression, affects the speaker's perceived certainty. Participants are placed in a listener's role: they are presented with utterances and asked to rate the speaker's certainty.

Degen et al. (2019) previously established that listeners ascribe varying degrees of certainty to a speaker depending on whether they use an evidential expression, as well as on the specific evidential expression they use. Importantly, all the evidential expressions they examined (English *must*, *might*, and *probably*, and German *muss* ('must'), *vermutlich* ('probably') and the discourse particle *wohl* (lit. 'well')) imply that the speaker has indirect evidence. As such, their results could be attributed to either extralinguistic reasoning about evidence directness, or to Gricean reasoning about competition with the bare form. By manipulating the directness of evidence, the present experiment can differentiate between these two accounts. In addition to utterances with evidential expressions implying evidence of varying directness, bare utterances are included as a baseline condition.

40 monolingual English speakers were recruited over the Prolific crowd-sourcing platform. Participants read the following brief story introducing the speaker and the discourse context:

(33) Your friend Taylor is at a party which you could not attend. The party is pretty fancy, but also crowded and noisy. Over the course of the party, Taylor is texting you about the people

at the party and what they are wearing. Each of the following statements is a text you receive from Taylor.

The speaker was described as the listener's friend to facilitate the assumption that she is cooperative. The party was described as crowded and noisy to facilitate the assumption that the speaker may be uncertain about the things she observes.

Participants then saw a total of 20 statements, each one describing a person's item of clothing using a single adjective. The critical items were 5 bare utterances with no evidential expression, and 5 utterances with the copulative perception verb *looks*. In addition, there were 10 filler items, 5 with the modal *might* and 5 with the phrase *I think*. Examples of the four kinds of statements are given below:

(34) a. Elliot's suit is beige (Bare)
b. Holly's dress looks purple (*looks*)
c. Mark's outfit might be crimson (*might* filler)
d. I think Tom's vest is green (*think* filler)

After each statement, participants were asked about the speaker's certainty about the corresponding bare proposition, e.g., "Is Taylor sure that Elliot's suit is beige?" and adjusted a slider with endpoints labeled "Absolutely sure" (coded as 100) and "Not sure at all" (coded as 0).

After the first statement they saw, and after 4 other statements, participants were additionally asked about the speaker's evidence for the proposition, e.g., "Why does Taylor think that Elliot's suit is beige?" Participants were given four potential sources of evidence to choose from: visual ("Taylor saw it"), haptic ("Taylor touched it"), reported ("Someone told Taylor about it"), or indirect inference ("Taylor has known Elliot for a long time"). These questions were included to ensure that participants paid attention to the speaker's choice of evidential expression.

Adjectives in the critical items were randomly selected from the list of adjectives examined by Lynott and Connell (2009), with the following constraints: adjectives were included only if they were familiar to all of Lynott and Connell's participants, had a frequency greater than 1 in the British National Corpus, and were not predicates of personal taste or aesthetic judgment, since such predicates allow a non-evidential reading of *looks* (see chapter 1, McNally & Stojanovic 2017, Poortvliet 2018). The 10 adjectives used in Experiment 1 and their mean visual strength

ratings (on a scale of 0 to 5) were: *purple* (5.00), *shiny* (4.95), *short* (4.95), *clean* (4.62), *striped* (4.52), *beige* (4.48), *bulky* (4.43), *loose* (4.14), *oily* (3.90), and *fuzzy* (3.67).

The experiment was created on www.qualtrics.com. Two lists were created, each containing 5 adjectives in bare utterances and 5 with the evidential expression *looks*, in addition to the 10 filler items. Participants were randomly assigned to one of the two lists, and items were presented to participants in pseudo-random order. The full list of materials is available online at https://osf.io/mvd9c/.

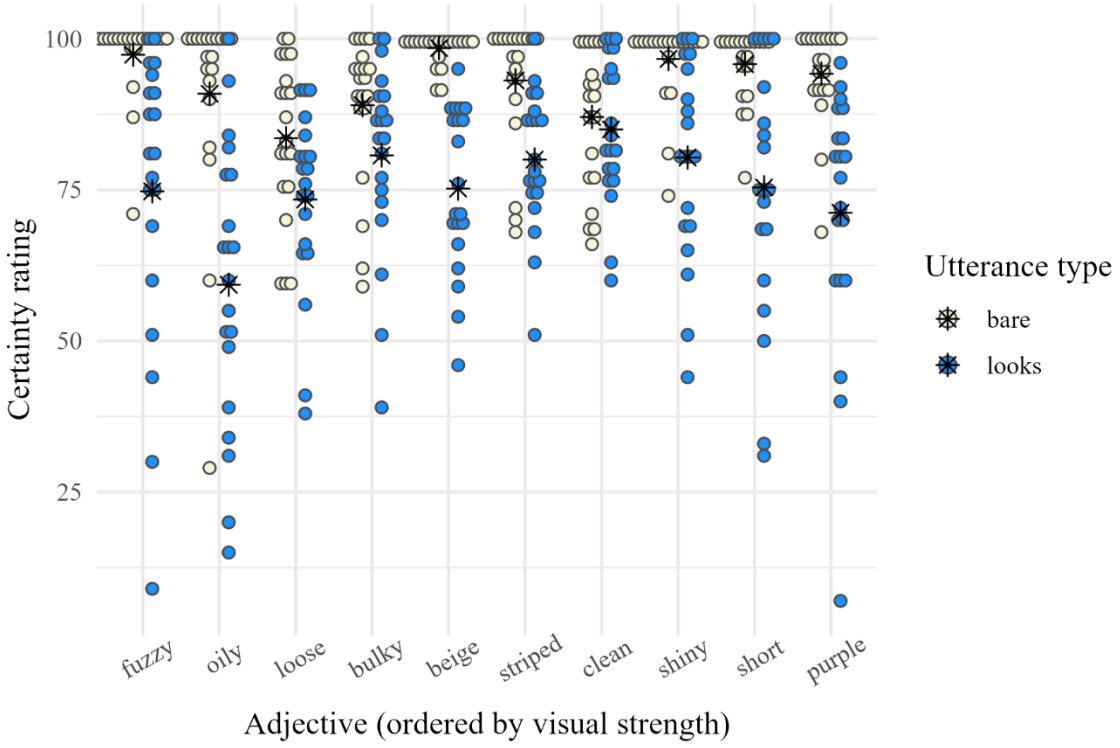


Figure 3. Certainty ratings in Experiment 1, grouped by adjective and utterance type. Asterisks represent mean ratings.

Certainty ratings grouped by utterance type and adjective are shown in Figure 3. Mean ratings for bare utterances are higher than for utterances with the evidential expression, across all adjectives. However, a visual inspection of the differences in means suggests that they are smaller for adjectives with medium visual strength (with the exception of *beige*). In other words, the evidential expression's uncertainty effect appears to be stronger for adjectives with either high or low visual

strength, hinting at a possible quadratic relation. I therefore incorporate a quadratic version of modality strength into our statistical model.

I conducted a mixed-effects regression predicting speaker certainty from a dummy-coded fixed effect of utterance type (bare as reference level), standardized modality strength, and standardized modality strength squared, as well as by-participant and by-adjective random intercepts. The data and the scripts used to analyze them are available online at https://osf.io/mvd9c/. Speaker certainty was rated as lower for utterances with the evidential expression than for bare utterances (β = -14.04, SE = 2.02, t = -6.96, $p < .0001$). The model yielded no simple effect for either modality strength (β = 1.92, SE = 1.63, t = 1.17, p = .264) or squared modality strength (β = 1.74, SE = 1.59, t = 1.09, p = .296), but the interaction between utterance type and squared modality strength was statistically significant: for utterances with the evidential expression, certainty was rated as lower with extreme values of modality strength (β = -3.04, SE = 1.53, t = -1.98, p = .048).

The results of the current experiment support the predictions of the Gricean account, in that the use of an evidential expression always conveys uncertainty relative to a bare utterance. The results could also be said to support the predictions of an extralinguistic account, after a fashion: the uncertainty effect of using an evidential expression depends on the directness of the evidence implied. However, contrary to the prediction that speaker certainty is proportional to evidence directness, the relation between them appears to be quadratic rather than linear. Specifically, the uncertainty effect is amplified for both relatively indirect evidence and maximally direct evidence. The latter effect cannot be explained by a purely extralinguistic account.

I take these results as supporting my hybrid account, which involves both Gricean reasoning and extralinguistic reasoning about evidence type. I predict that the increased uncertainty associated with utterances implying maximally direct evidence is the result of an extra inference, that the evidence is compromised, for instance by poor perceptibility conditions. This prediction is tested in Experiment 2.

2.3.2 Experiment 2: Production

This experiment investigates how evidence directness and perceptibility conditions affect the choice between using a bare utterance and using an evidential expression. Participants are here

placed in a speaker's role: they are presented with a context and asked to choose between possible utterances.

Pogue and Tanenhaus (2018) and Degen et al. (2019) have both explored choice of utterance as a function of the evidence available to the speaker. Pogue and Tanenhaus presented participants with visual evidence in the form of images, and manipulated the completeness of the images and the amount of time participants had to view them. Degen et al. presented participants with textual descriptions of various types of evidence: perceptual, reported, and inferential.

Essentially, Pogue and Tanenhaus kept the type of evidence constant and manipulated its reliability, while Degen et al. did the opposite, manipulating evidence type and not reliability. Hence, in both studies, the evidence presented to participants could be ranked on a single scale of evidence strength. Both studies found that speakers were more likely to use an evidential expression with weak evidence, and more likely to use a bare utterance with strong evidence.

The present experiment explores a more complex conception of evidence strength, comprising both directness and reliability. As in Experiment 1, evidence directness is based on Lynott and Connell's (2009) sensory modality strength norms for adjectives. Unlike Experiment 1, three different sensory modalities are included: visual, auditory, and haptic. Evidence reliability is manipulated with textual descriptions of good and poor perceptibility conditions.

33 monolingual English speakers were recruited over the Prolific crowd-sourcing platform. Participants were presented with 12 brief texts describing situations and asked to choose between two utterances to use in each situation. Each text described the speaker standing outside a room which they cannot enter, and about which they have only a single source of information. After reading the text, participants were asked what they would say to a friend who lacked access to the same information, should that friend ask them whether the room had a certain property.

The 6 critical items were situations in which the participants' source of information was visual (a window they could look through), auditory (a door they could listen at), or tactile (a gap under the door they could reach through). There were also 6 filler items in which the participants' source of information was olfactory, linguistic, or mixed. In each situation, the source of information was described as either good or poor. Two examples of situations, one with good visual evidence and one with poor auditory evidence, are the following:

(35) a. Imagine that you are standing outside a room. You can't hear anything inside, but there is a window that you can look through. The window is perfectly clear, so you can see what it's like in the room very well. (Good visual evidence)

b. Imagine that you are standing outside a room. You can't see inside, but you can listen at the door. However, the door is very thick, so it's difficult to hear what it's like in the room. (Poor auditory evidence)

In the critical items, participants were asked to choose between a bare utterance and an utterance with a copulative perception verb matching the source of information: *looks*, *sounds*, or *feels*. In the filler items, participants were asked to choose between two utterances with two different copulative perception verbs.

Again, adjectives in the critical items were selected from the list of adjectives examined by Lynott and Connell (2009). Adjectives were included only if they were familiar to all of Lynott and Connell's participants, had a frequency greater than 1 in the British National Corpus, and were not predicates of personal taste or aesthetic judgment. For Experiment 2, I selected "weather" predicates, which could occur in impersonal constructions, e.g., *it's hot in there*. The 6 adjectives used in Experiment 2 and their mean modality strength ratings for the three studied modalities (on a scale of 0 to 5) are presented in Table 10.

Table 10. Adjectives used in Experiment 2 and their mean modality strength ratings (Lynott & Connell 2009).

Adjective	Mean visual strength	Mean auditory strength	Mean haptic strength
bright	5.00	0.14	0.19
crowded	4.62	3.71	2.29
wet	4.33	1.86	4.67
hot	3.33	1.05	4.86
humid	1.76	0.24	3.29
noisy	1.67	4.95	0.29

The experiment was created in Google Forms.[23] Six lists were created, each containing two situations with each source of information, one good and one poor. Participants were randomly assigned to one of the lists, and items were presented to participants in pseudo-random order. The order in which utterances were presented as choices to participants was alternated between critical items. The full list of materials is available online at https://osf.io/mvd9c/.

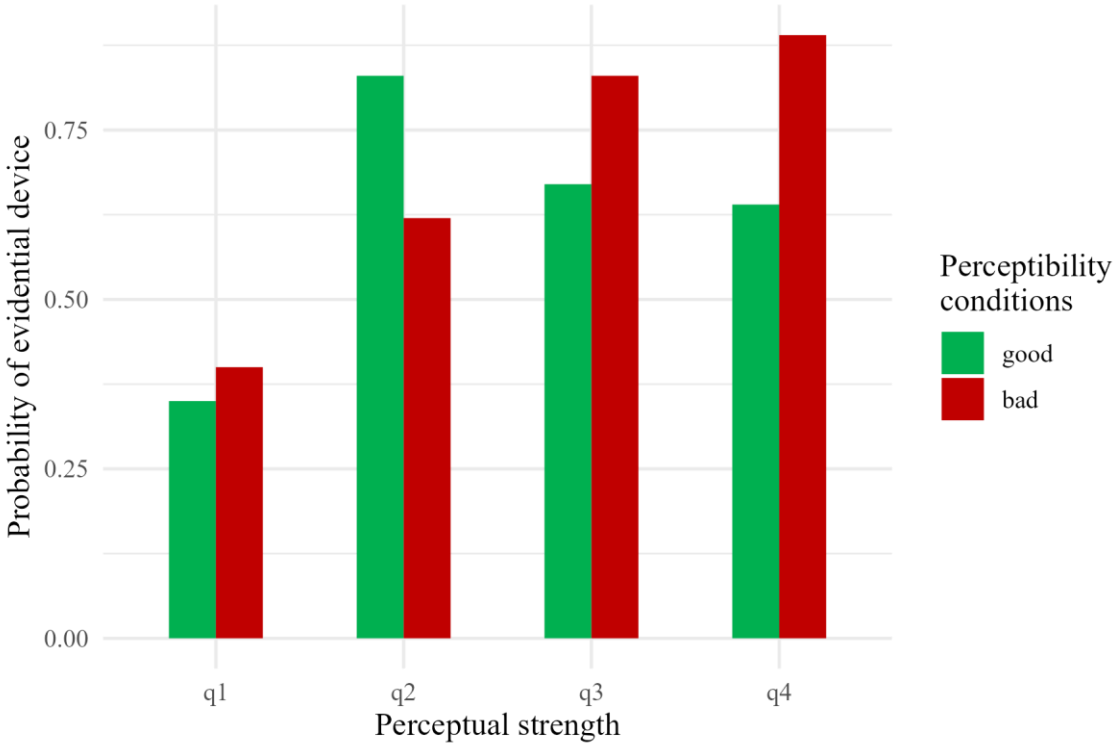


Figure 4. Probability of participants choosing to use an evidential expression (and not a bare utterance) in Experiment 2, grouped by perceptibility conditions and modality strength quadrant.

The probability of using an evidential expression, by perceptibility conditions and modality strength quadrants, is presented in Figure 4. I conducted a mixed-effects logistic regression predicting the likelihood of using an evidential expression from a dummy-coded fixed effect of perceptibility (good as reference level), a sum-coded fixed effect of sensory modality (visual as

[23] The reasons for creating the two experiments on two different platforms are technical. Experiment 1 was created in www.qualtrics.com because that platform allows the creation of questions which participants respond to by adjusting a slider. Experiment 2 was created in Google Forms instead because of its friendlier interface for exporting results.

reference level), standardized modality strength, and standardized modality strength squared, as well as by-participant and by-adjective random intercepts. The data and the scripts used to analyze them are available online at https://osf.io/mvd9c/.

The model yielded no main effect for perceptibility (β = -0.49, *SE* = 0.62, p = .429) and no simple effects for modality strength (β = 0.22, *SE* = 0.25, p = .388) or sensory modality (auditory: β = 0.16, *SE* = 0.28, p = .583; haptic: β = 0.02, *SE* = 0.26, p = .952). In good perceptibility conditions, the likelihood of using an evidential expression decreased for extreme values of modality strength (β = -1.01, *SE* = 0.34, p = .003). Conversely, in poor perceptibility conditions, the likelihood of using an evidential expression increased linearly with modality strength (β = 0.89, *SE* = 0.37, p = .016).

To compare the results of the Experiment 2 to those obtained in Experiment 1, the notion of evidence strength must be reconstructed from its two material components: directness and reliability. Evidence is strongest when it is both direct, represented here by high modality strength, and reliable, represented here by good perceptibility. With strong evidence, we expect high speaker certainty, and consequently more bare utterances. This prediction is borne out. As evidence becomes weaker, due to either indirectness or unreliability, we expect more uncertainty, and consequently more evidential expressions. This prediction too is borne out.

Where my account's predictions seem to fail is with the very weakest evidence, where we would expect the greatest degree of uncertainty, but find minimal use of evidential expressions (see the lower end of the modality strength scale in Figure 4). I attribute this to a limitation of the experimental design: participants were forced to choose between a bare utterance and an utterance with an evidential expression, in a context where neither option is apt. In general, answering a question based on very weak evidence (e.g., answering whether a room is noisy based only on haptic evidence) is not an exemplar of cooperative pragmatic behavior. Many speakers in such circumstances would prefer to admit ignorance or say nothing, but these options were not available to my participants.

I can think of two reasons why my participants were particularly averse to using an evidential expression when the evidence indicated by it was very weak. First, it's possible that a minimum threshold of evidence strength is "hardcoded" into the semantics of copulative perception verbs. In other words, utterances such as *It feels noisy in there* may be grammatically unacceptable, in

addition to pragmatically uncooperative. Conversely, any constraint on the evidence supporting bare utterances is solely pragmatic, e.g., Grice's (1989) maxim of quality. A second possibility is that a minimum of evidence strength is required for an utterance with an evidential expression to be interpreted as addressing the question under discussion (Roberts 2004), otherwise it comes off as a non sequitur. Again, this is not a pitfall that bare utterances can fall into. These two tentative explanations could potentially be teased apart experimentally, but I leave this to future research.

2.4 Computational model

I formalize my account within the Rational Speech Act (henceforth RSA) framework (Frank & Goodman 2012, Goodman & Frank 2016). The framework models cooperative pragmatic behavior as recursive probabilistic reasoning between rational agents. The simplest of these agents is a hypothetical "literal listener", L_0, which interprets utterances according to their literal meaning. The literal listener is modeled as a function which takes an utterance as input, and outputs a distribution of probabilities over states in which the utterance is true (potentially weighted by states' prior probabilities, representing world knowledge). Next is the pragmatic speaker, S_1, which chooses an utterance with the literal listener in mind, seeking to maximize informativity – the probability of L_0 inferring the correct intended meaning – while minimizing their own effort. The pragmatic speaker is modeled as a function which takes an intended meaning (typically an observed state of the world) as input, and outputs a distribution of probabilities over utterances. Finally, the pragmatic listener, L_1, uses Bayes' rule to "invert" the pragmatic speaker's utterance probabilities, in order to recover the most likely meaning intended by S_1.

To demonstrate the nuts and bolts of the framework, consider the following simple setup (adapted from Goodman & Stuhlmüller 2013): there are two apples, each of which may have gone bad. Hence, the set of possible states of the world W is given in (36). Assume that the speaker can only use the set of utterances U in (37). Taking the traditional semantics of *some* as an existential quantifier, the utterance *Some of the apples are bad* would be literally true if at least 1 apple is bad.

(36) W = {0 bad apples, 1 bad apple, 2 bad apples}

(37) U = {*Some of the apples are bad, All of the apples are bad, None of the apples are bad*}

With this setup, the literal listener L_0, upon hearing *Some of the apples are bad*, will ascribe equal probabilities P to the two states in which that utterance is true (assuming it had no prior knowledge of how likely it is for an apple to go bad); this is shown in (38a). Meanwhile, hearing *All of the apples are bad* (or *None of the apples are bad*) leaves only one possible state, as shown in (38b):

(38) a. L_0(*Some of the apples are bad*) → P(1 bad apple) = 50%;

P(2 bad apples) = 50%

b. L_0(*All of the apples are bad*) → P(2 bad apples) = 100%

The pragmatic speaker S_1, upon observing 2 bad apples, has two literally true utterances to choose from, but one of them would be more informative to the literal listener: using *all* guarantees that L_0 will arrive at the correct state of the world, whereas using *some* leaves it to chance. S_1 is more likely to choose a more informative utterance (using a mathematical procedure to assign an informativity score to each utterance, and convert informativity scores into probabilities). For the present setup, the qualitative result is shown in (39):

(39) S_1(2 bad apples) → P(*All of the apples are bad*) > 50%;

P(*Some of the apples are bad*) < 50%

The pragmatic listener L_1 contains its own representation of S_1. Upon hearing *Some of the apples are bad*, it computes, just like the literal listener, that there are in principle two states in which that utterance is true. However, it also computes that, if S_1 were to observe 2 bad apples, it would be more likely to use the more informative *all*. It thus "infers" that, since S_1 used *some*, it is more likely to have observed 1 bad apple. Consequently, L_1 ascribes a higher probability to that state; see (40). Thus, L_1 effectively derives the scalar implicature from *some* to *some but not all*, using traditional semantics and general principles of rationality.

(40) L_1(*Some of the apples are bad*) → P(1 bad apple) > 50%;

P(2 bad apples) < 50%

The RSA framework has been extended to capture a wide array of pragmatic phenomena, including scalar implicatures (Goodman & Stuhlmüller 2013, Degen & Goodman 2014), metaphor understanding (Kao et al. 2014a), hyperbole (Kao et al. 2014b), and inferred speaker uncertainty

(Degen et al. 2015). For my own model, meant to capture the discourse effects of evidential expressions, I adopt some earlier extensions and add several novel ones. Underlying my own extensions are some nontrivial assumptions regarding beliefs and communication, spelled out below.

I assume that people's beliefs about the state of the world are based on the evidence available to them, specifically on how direct and how reliable that evidence is. In communicating about the state of the world, speakers can choose to either directly describe what they believe the world to be like, using a bare utterance, or else to describe what evidence they have for a belief, using an evidential expression. Speakers and listeners alike know that evidence is not always perfect, but assume that evidence tends to match reality more often than not (otherwise it wouldn't count as evidence). A speaker does not have to be perfectly certain to use a bare utterance, but using an evidential expression can nevertheless be the "safer" choice, in the sense that it avoids committing to something that could turn out to be false. On the other hand, the unmarked bare utterance is often more informative. Speakers also always have the choice to say nothing at all.

My model is designed to reflect the setup of Experiment 2, described in section 2.3. The space of possible utterances *U* thus includes bare utterances, utterances with each of the evidential expressions *looks*, *sounds*, and *feels*, and a "null utterance", which represents saying nothing. For any property *Q* used in Experiment 2, the utterance space is:

(41) U_Q = {*It's Q*, *It looks Q*, *It sounds Q*, *It feels Q*, NULL}

I assume simplified semantics for these utterances. A bare utterance *It's Q* is true iff $Q(x)$ holds of the relevant entity *x*.[24] For any evidential expression *EVID* which implies evidence of type ε, the utterance *It EVIDS Q* is true iff there is evidence of type ε that $Q(x)$ holds of the relevant entity *x*.[25] The null utterance is always true. To capture how evidence does not always attest to actual fact, and how different types of evidence may be at odds, I represent states of the world as *n*-tuples. The first element in the *n*-tuple is the "actual" state of the world, to which bare utterances refer. Subsequent elements are "evidence states", each of which represents whether or not there is

[24] This abstracts away from degree semantics, e.g., just how bright must *x* be for *x is bright* to be true (see Lassiter & Goodman 2013 for an RSA model tackling this issue).

[25] Although simplistic, this semantics is compatible with recent analyses of evidentials, and evidential uses of copulative perception verbs, using possible world semantics (Muñoz 2019, Faller 2020).

evidence of a particular type for the actual state. Thus, for any proposition q and evidence for q of types ε… ε', the set of possible states is:

(42) $W_q = \{q,\neg q\} \times \{\varepsilon_q, \neg\varepsilon_q\} \times \ldots \times \{\varepsilon'_q, \neg\varepsilon'_q\}$

This representation captures the incontrovertible fact that evidence does not always match reality. However, I also want to capture the commonsense intuition that evidence does tend to match reality more often than not (otherwise it wouldn't be considered evidence). I incorporate this intuition into listeners' prior assumptions about the state of the world: states in which evidence matches reality are a priori more likely than states in which evidence does not match reality. Mathematically, this is represented with the parameter $E > 1$, the "evidence coefficient". For any proposition q, the prior probability of states in which q is true (i.e., the actual state is q) is greater by a factor of E for each type of evidence ε there is for q (i.e., each evidence state where ε_q is true):

(43) $P(<q, \varepsilon_q, \varepsilon'_q>) = E \cdot P(<q,\neg\varepsilon_q, \varepsilon'_q >) = E \cdot P(<q, \varepsilon_q, \neg\varepsilon'_q >) = E^2 \cdot P(<q, \neg\varepsilon_q, \neg\varepsilon'_q >)$

This is what allows utterances with evidential expressions to address questions about the actual state of the world (i.e., for the evidenced proposition to be their at-issue content; see Murray 2017, Faller 2019). The literal listener, upon observing an utterance with an evidential device, assigns probabilities to states in which that utterance is true, including to states in which the evidenced proposition is false. The evidence coefficient ensures that states in which the evidenced proposition is true would be assigned a higher total probability than states in which the evidenced proposition is false.

To see this at work, consider the set of possible states in (44a). The evidence coefficient ensures that some of these states have higher prior probabilities than others; see (44b). The literal listener L_0, upon observing the utterance *It looks bright*, assigns probabilities to states in which that utterance is true, i.e., states in which there is visual evidence that the relevant entity is bright. There are two such states, and the one in which the relevant entity really is bright has a higher prior probability (exactly how much higher depends on the value of E); see (45).

(44) a. $W = \{<bright(x), \text{VISUAL}_{bright(x)}>, <\neg bright(x), \text{VISUAL}_{bright(x)}>,$

$<bright(x), \neg\,\text{VISUAL}_{bright(x)}>, <\neg bright(x), \neg\,\text{VISUAL}_{bright(x)}>\}$

b. $P(\langle bright(x), \text{VISUAL}_{bright(x)}\rangle) = E \cdot P(\langle \neg bright(x), \text{VISUAL}_{bright(x)}\rangle) =$

$E \cdot P(\langle bright(x), \neg \text{VISUAL}_{bright(x)}\rangle) = P(\langle \neg bright(x), \neg \text{VISUAL}_{bright(x)}\rangle)$

(45) $P_{L0}(\textit{It looks bright}) \rightarrow P(\langle bright(x), \text{VISUAL}_{bright(x)}\rangle) > 50\%$;

$P(\langle \neg bright(x), \text{VISUAL}_{bright(x)}\rangle) < 50\%$

For present purposes, we are only interested in the listener's beliefs about the actual state, under the assumption that by default, the at-issue content of an utterance with an evidential expression is the evidenced proposition (Murray 2017, Faller 2019, cf. Korotkova 2020).[26] Thus, we define the literal listener function L_0 as follows: upon observing an utterance u with at-issue content q, L_0 sums the prior probabilities of states $w_{u,q}$ in which both u and q are true, in order to compute the probability of q. This is formally represented in (46).

(46) $P_{L0}(q|u) \propto \sum P(w_{u,q)}$

In a basic RSA model, speakers are assumed to be perfectly certain about the state of the world they wish to communicate. As a corollary, speakers are also assumed to never produce utterances whose truth they are not certain of. Both of these assumptions are too strong for present purposes. To capture speaker uncertainty, I represent speaker beliefs as a probability distribution over possible states (see Goodman & Stuhlmüller 2013, Scontras et al. 2018, for a similar approach). Specifically, I assume that speakers' beliefs are based on the directness and reliability of the evidence available to them (see section 2.2). Mathematically, DIR(ectness) is a function of a proposition q and evidence for q of type ε, and REL(iability) is a function of evidence type ε and context c; see (47a-b). The speaker's belief regarding q, upon observing evidence for q of type ε in context c, is computed as in (48).

(47) a. $\text{DIR}(q,\varepsilon_q) \in (0,1)$

b. $\text{REL}(\varepsilon,c) \in (0,1)$

[26] There may be contexts in which the at-issue content of an utterance with an evidential expression is not the evidenced proposition, but the presence or absence of evidence, e.g., in response to the question *What makes you think that the dress is blue?* The present model could potentially be extended to capture such cases, by taking inspiration from Kao et al. (2014b) RSA model, in which the pragmatic listener infers the question under discussion.

(48) $P_{S1}(q|\varepsilon_q,c) = 0.5 + 0.5 \cdot \text{DIR}(q,\varepsilon_q) \cdot \text{REL}(\varepsilon,c)$

With maximally direct and reliable evidence for a proposition, the speaker's belief in it approaches certainty (1), but when evidence is either very indirect or very unreliable (or both), belief approaches agnosticism (0.5). This belief function plays a role in the pragmatic speaker's choice of utterance. Namely, the pragmatic speaker function S_1 seeks to minimize the divergence between its own belief state, and the expected belief state of the literal listener. The idea is that a speaker who is very certain, having observed direct and reliable evidence, will want their addressee to be very certain as well. Conversely, if the speaker is not quite certain, having observed less direct or less reliable evidence, they will want the addressee to maintain some doubt as well, to minimize the risks of losing face and trust.

Mathematically, S_1 assigns (negative) informativity scores to each utterance u based on the Kullback-Leibler divergence between its own belief state, $P_{S1}(\cdot|\varepsilon_q,c)$, and the belief state of the literal listener observing u, $P_{L0}(\cdot|u)$. These scores are then multiplied by the "optimality" parameter $\alpha > 0$, which represents how committed the speaker is to choosing the most informative utterance, and finally exponentially transformed into probabilities:

(49) $P_{S1}(u|\varepsilon_q,c) \propto exp(-D_{\text{KL}}(P_{S1}(\cdot|\varepsilon_q,c) \,||\, P_{L0}(\cdot|u)) \cdot \alpha)$

To illustrate how S_1 works, consider a case where the speaker has visual evidence, obtained in a context with good visibility, that an entity is bright. Since the evidence is both direct and reliable, the speaker will have a strong belief in the relevant proposition. The speaker then has to choose between two true utterances: *It's bright* and *It looks bright*. To do so, it calculates which of the two utterances will bring the literal listener to a belief state closer to the speaker's own. Because *It looks bright* leaves open the possibility that the evidenced proposition is false, the answer is likely going to be the bare utterance *It's bright*.

The last component of my model is the pragmatic listener. The pragmatic listener function L_1 performs a joint inference, "reasoning" not only about the actual state, but also about the evidence available to the speaker. Upon observing an utterance, L_1 outputs three probability distributions: over the actual state, over the type of evidence available to the speaker, and over the conditions under which the evidence was obtained. It does so by using Bayes' rule to "reverse engineer" the

circumstances most likely to have led the speaker to produce the observed utterance, re-weighted by the listener's prior beliefs:

(50) $P_{L1}(q,\varepsilon_q,c) \propto P_{S1}(u|\varepsilon_q,c) \cdot \sum P(w_{u,q})$

To illustrate how L_1 works, assume it hears the utterance *It looks bright*. It then attempts to recover the circumstances which led the pragmatic speaker to choose that utterance. If the speaker had a very strong belief, it would've chosen to use the bare utterance *It's bright*, following the logic outlined above. If the speaker had auditory or haptic evidence, it would've chosen to use *sounds* or *feels*, respectively. Therefore, the speaker must have visual evidence that leads to a weak belief. Since visual evidence for brightness is very direct, the only reason for the speaker to have a weak belief is that the evidence is unreliable. Thus, the pragmatic listener will likely infer that the speaker is uncertain about the proposition, has visual evidence for it, and that evidence was obtained in a context that makes it unreliable.

To have the model produce predictions, the various parameters must be given numerical values. For the evidence coefficient E, I use the average certainty rating obtained in Experiment 1 for utterances with evidential expressions, divided by agnosticism: $E = 0.75/0.5 = 1.5$. To generate directness values between 0 and 1 for pairs of propositions and evidence types, I divide Lynott & Connell (2009)'s modality strength norms by 5. For simplicity's sake, I assume just two possible values for reliability, corresponding to the good and poor perceptibility conditions in Experiment 2. For any evidence type ε, $\text{REL}(\varepsilon,\text{good}) = 0.9$ and $\text{REL}(\varepsilon,\text{poor}) = 0.1$. Lastly, I set $\alpha = 10$. The model's predictions with these parameter values are shown in Figures 5 and 6 below for the pragmatic listener and the pragmatic speaker, respectively.

The model is able to capture some of the qualitative findings of the empirical experiments reported in section 2.3. On the listener's side, the pragmatic listener function consistently associates higher certainty with bare utterances compared to utterances with an evidential expression. Moreover, when given utterances with evidential expression as input, the pragmatic listener produces a quadratic relation between modality strength and certainty: certainty is lower both for low evidence directness and for maximally high directness.

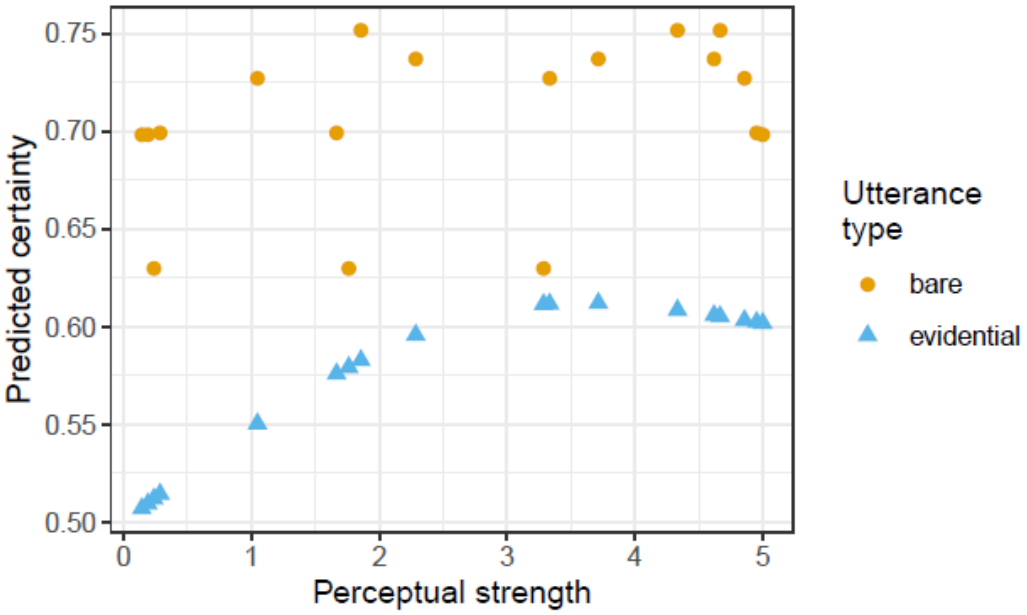


Figure 5. The computational model's predictions for the pragmatic listener function's inferences regarding the speaker's certainty in an at-issue proposition, depending on the speaker's utterance choice.

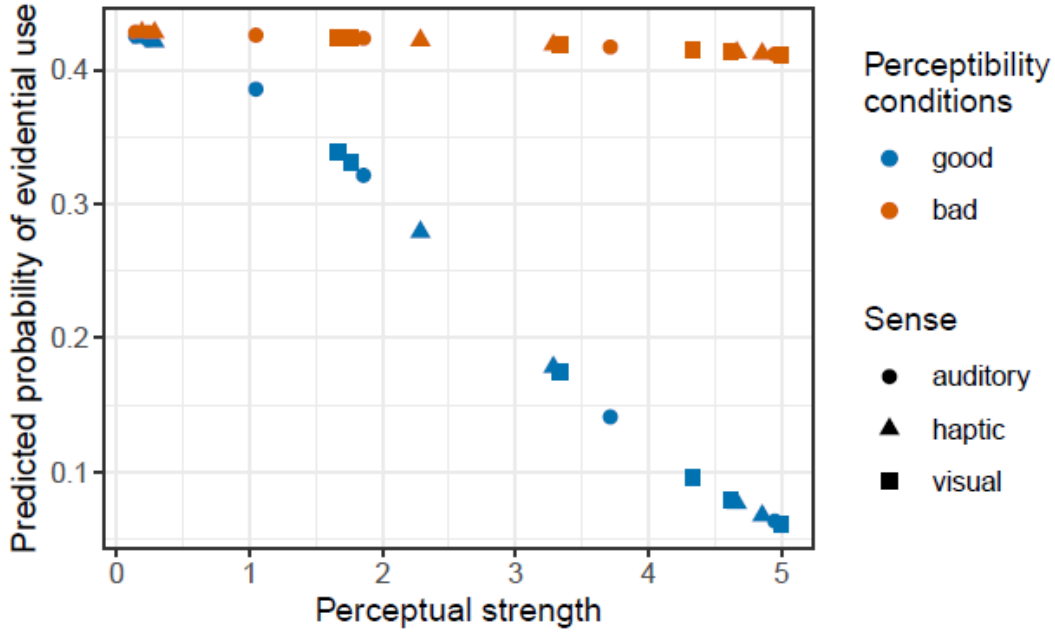


Figure 6. The computational model's predictions for the pragmatic speaker function's utterance choice, depending on the directness (perceptual strength) and reliability (perceptibility conditions) of evidence available to them.

On the speaker's side, the pragmatic speaker's probability of using an evidential expression depends on the interaction between modality strength and perceptibility conditions. In good conditions, the use of evidential expressions decreases with modality strength, but in poor conditions, it does not. All in all, the model is able to derive both the uncertainty inference associated with evidential expressions, and the doubt-or-denial condition.

Unfortunately, the empirical data collected in the experiments are not suited for a quantitative comparison with the model's predictions. This is due to the limitations of the experimental designs: Experiment 1 included only a single sensory modality and a relatively narrow range of modality strength ratings, and Experiment 2 forced participants to choose between two potentially inappropriate utterances. A more rigorous testing of the model will require extended designs for both.

2.5 Conclusions

In this chapter I have shown that the uncertainty inference associated with evidential devices is considerably more complex than predicted by previous accounts. Experiment 1 revealed that the use of an evidential expression conveyed uncertainty even when implying maximally direct evidence. Moreover, this uncertainty inference was enhanced for maximally direct evidence. These findings challenge extralinguistic accounts, which predict that inferred speaker certainty is directly proportional to the directness of evidence available to the speaker, as well as purely Gricean accounts, which predict no effect for evidence directness.

It is my claim that only a hybrid account, involving both extralinguistic and Gricean reasoning, can explain the results of Experiment 1. My own account does so by introducing a notion of evidence strength comprised of two distinct elements: directness and reliability. My account's predictions were partly borne out in Experiment 2, which revealed that evidential expressions were most likely to be used either with high directness and low reliability, or with midrange directness and high reliability.

I believe that this research constitutes an important step toward a better understanding of evidential expressions and their discourse effects. More generally, it contributes to the study of how beliefs are formed, assessed, and communicated. In addition, this work illustrates the intricacies involved in this sorely understudied domain, as well as, I hope, the advancements that could be made in it using novel experimental and computational tools.

3 Synesthetic metaphors

3.1 Introduction

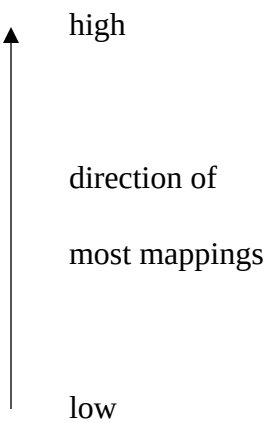
A synesthetic metaphor is a description of a percept in terms of a different sensory modality. For example, the phrase *soft brightness* is a description of a visual percept in tactile terms. As their name implies, synesthetic metaphors are instances of metaphorical mapping, from a source domain to a target domain. In the case of *soft brightness*, these are the domains of touch and sight, respectively.

Perhaps the most widely discussed issue in the literature on synesthetic metaphor is that of directional preferences. The idea behind this is that certain types of mappings are somehow better than their opposites, e.g., that touch-to-sound is better than sound-to-touch. Ullmann (1945, 1957) famously proposed that directional preferences conform to a hierarchy of the senses, such that mappings "upward" on the hierarchy, i.e., from low modalities (touch, taste) to high modalities (sound, sight), are better than their opposite, "downward" mappings. Thus, phrases like *soft brightness* (touch-to-sight), *tasty noise* (taste-to-sound), and *chilled scent* (touch-to-smell), are expected to be more frequent in discourse, judged as more natural, recalled better, etc., than their opposites: *bright softness* (sight-to-touch), *noisy taste* (sound-to-taste), and *scented chill* (smell-to-touch).

Ullmann (1945, 1957) pioneered the empirical study of synesthetic metaphors with a quantitative analysis of 19th century English, French and Hungarian literary corpora. He collected several hundred synesthetic metaphors in various grammatical forms, disregarding what he judged to be "stale" or conventionalized metaphors, and annotated their target and source domains.

From his findings, Ullmann draws three principal generalizations. First, the majority of transfers (i.e., mappings) are directed from lower toward higher levels of the sensorium. This generalization relies on the assumption, adopted from classical philosophy, that the senses are ordered hierarchically as in (51) below. Ullmann's second generalization is that most of the transfers are taken from the sphere (i.e., domain) of touch, and the third is that most of the transfers are directed toward the sphere of sound. A corollary of the two latter generalizations is that the single most frequent type of transfer is from touch to sound.

(51)



Ullmann explicitly raises the possibility that these generalizations represent a semantic law, although there are exceptions to this proposed law even in his own data, namely in mappings between sight and sound: downward mappings from sight to sound are more frequent in his corpora than the opposite upward mappings. As a possible explanation for this exception, Ullmann remarks that sound has fewer words associated with it than sight, and hence is more likely to "recruit" descriptors from other domains.

Following Ullmann's work, a great deal of research on synesthetic metaphors has been devoted to corroborating, extending, refining, or explaining some or all of Ullmann's generalizations regarding directional preferences (Dombi 1974, Williams 1976, Day 1996, Wise 1997, Shen & Cohen 1998, Yu 2003, Werning et al. 2006, Shen & Eisenman 2008, Shen & Gil 2008, Shen & Gadir 2009, Shinohara & Nakayama 2011). These include corpus studies reproducing Ullmann's generalizations in various languages and genres, including Hungarian poetry (Dombi 1974), English and German literature (Day 1996), Hebrew poetry (Shen 1997), and Chinese literature (Yu 2003).

Williams' (1976) diachronic study of English and Japanese sensory adjectives extends Ullmann's generalizations from novel metaphorical mappings to fully conventional ones. Later works likewise extend Ullmann's generalizations, from analyses of naturally occurring data, to interpretation of novel experimental materials (e.g., Shen 1997, Shen & Eisenman 2008, Shen & Gil 2008, Shen & Gadir 2009, Shinohara & Nakayama 2011).

This entire research paradigm, however, has recently come under criticism. Winter (2016, 2019a, 2019b; see also Ronga et al. 2012) critiques many of its underlying assumptions and methodological practices, as well as the theorizing behind it. First, many of the studies following

Ullmann assume a clear delineation of sensory domains. Although they differ in the ways they "carve up" the sensory conceptual space, e.g., whether they separate touch from heat (Ullmann 1945) or color from dimension (Williams 1976), they generally agree that human sensory experience can be delineated into 5 to 8 independent domains. As Winter notes, the various delineations often reflect a particular researcher's cultural framework rather than any established psychophysical theory.

Second, studies within this paradigm tend to adopt a categorical approach to sensory words, whereby each such word is taken to evoke a single sensory domain, and all associations between sensory words and domains are considered equal. This assumption contrasts with experimental and corpus evidence that words may be associated with several sensory domains, and to different degrees (Lynott & Connell 2009, 2013, Winter 2016). Moreover, it is the researchers themselves that usually code each sensory word as associated with one domain or another, instead of relying on more reproducible methods. For example, Shen & Gadir (2009) use Hebrew words they translate as *honey* and *form* to evoke taste and sight, respectively, in their experiment. Yet the participants in Lynott & Connell's (2013) study rated English *honey* as strongly experienced through sight ($M = 4.12$, $SD = 0.93$) and smell ($M = 3.76$, $SD = 1.15$) in addition to taste ($M = 4.76$, $SD = 0.56$), and in fact as more strongly experienced through sight than *form* ($M = 3.24$, $SD = 1.75$).

Similar to the above is the assumption that sensory words are always used to evoke their associated sensory domains, rather than with other conventional, non-sensory meanings. Winter (2019b) argues that many of the examples discussed in the literature on synesthetic metaphors are conventionalized, to the extent that they might not be considered metaphorical at all. For example, English *sweet* could be argued to have a fully conventional affective meaning, which no longer depends on its meaning as a taste word. Hence, phrases like *sweet melody* might be interpreted as straightforward affective evaluations, rather than metaphorical mappings across sensory domains. Even when researchers address the difference between conventional and novel metaphors, as when Ullmann (1945, 1957) sets aside "stale" metaphors, they tend to rely on their own judgments in doing so.

Finally, Winter (2016) critiques the theoretical accounts advanced in many of these studies, particularly the way they draw causal conclusions from correlational data. Various factors have

been proposed in the literature as potential causes for directional preferences, but only a few of these factors have been tested directly. Such testing requires either an experiment, where one factor is manipulated while other, potentially confounding factors, are controlled for; or careful statistical analyses of corpus data, revealing whether one or more factors reliably predict the occurrence of synesthetic metaphors.

Broadly speaking, the factors previously proposed as contributing to directional preferences fall into two camps (cf. Winter, 2019a). The first consists of perceptual factors, i.e., properties of the sensory modalities themselves, which are taken to have a direct effect on directional preferences. For example, the preference for *soft brightness* over *bright softness* might be due to differences between how we perceive light as opposed to how we perceive texture, e.g., that texture perception is somehow more embodied than light perception. Crucially, the effect of perceptual factors does not depend on the word choice in a given synesthetic metaphor, but only on the modalities involved.

The second camp consists of lexical factors, i.e., properties of the words associated with the different sensory modalities. For example, the preference for *soft brightness* over *bright softness* might be due to differences between the words *soft* and *bright*, e.g., that *soft* is more frequent or more affectively loaded than *bright*. Such lexical factors might still ultimately be traced back to perceptual factors, in that the properties of each sensory modality influence the makeup of its associated lexical field. Thus, the properties of *soft* and *bright* might be typical to touch-words and sight-words, respectively, because of differences between how we perceive light as opposed to how we perceive texture. However, this is an indirect effect of perceptual factors on directional preferences, contingent on the particular word choice in a given synesthetic metaphor.

An example of a perceptual factor is degree of embodiment, invoked in a speculative account by Shen and colleagues (Shen 1997, Shen & Eisenman 2008, Shen & Gadir 2009). They propose that touch and taste, as the only modalities that require direct contact between perceiver and stimulus, are more embodied than smell, sound and sight, and are therefore more cognitively accessible. This account subsumes directional preferences in synesthetic metaphors under the general principles of conceptual metaphor theory (Lakoff & Johnson 1980, 1999), which posits that metaphorical mappings generally occur from more accessible to less accessible domains. Degree of embodiment, defined in this way, is a perceptual factor, not a lexical factor: whether or not a

stimulus is perceived via direct contact depends on the sensory modality involved, not on the words used to describe the stimulus. To my knowledge, the effect of degree of embodiment on directional preferences has never been tested directly.

Shibuya et al. (2007) attempt to explain directional preferences specifically between touch and sight, by appealing to a notion of sensory association. They propose that associations between the senses, which allow for interpretable synesthetic mappings, are based on co-occurrences between sensory stimuli in daily experience. The relative frequency of the co-occurrence determines the strength of the association. For example, tactile stimuli almost always co-occur with visual stimuli, whereas only a minority of visual stimuli co-occur with tactile stimuli. Therefore, the association of touch with sight is stronger than the association of sight with touch, making mappings from touch to sight more interpretable than their opposites. Sensory association is another perceptual factor, since the co-occurrence of stimuli does not depend on the words used to describe the stimuli. Like embodiment, the effect of sensory association on directional preferences has not been tested directly in previous literature.

Popova (2005) as well as Petersen et al. (2007) discuss the notion of gradability in this connection, particularly antonymic, unbounded gradability, which is a lexical factor. They propose that metaphorical mapping of antonymic, unbounded gradable features is more natural than that of other features, because the former can be mapped onto abstract, modality-general scales. For example, softness is antonymic (the opposite of *soft* is *hard*) and unbounded (there's no maximal degree of either softness or hardness). It can therefore be mapped onto an abstract scale such as intensity or affectivity, allowing *soft* in *soft brightness* to be naturally reinterpreted as, say *faint* or *pleasing*. In contrast, redness is not antonymic (there's no unique opposite to *red*) and silence is bounded (one thing cannot be literally more silent than another). Therefore, *red* and *silent* cannot be mapped onto those same abstract scales, and don't receive a natural reinterpretation.

Popova (2005) further argues that touch (and to a lesser extent, taste) is associated with this kind of gradability more so than sound or sight, which explains why it makes a better source domain. Nevertheless, gradability is a lexical factor: a given percept, in any sensory modality, can be described using either gradable or non-gradable features, e.g., *quiet* (unbounded) vs. *silent* (bounded), or *bright* (antonymic) vs. *white* (non-antonymic). Petersen et al. (2007) present experimental evidence for the effect of gradability on directional preferences. They show that in

German synesthetic metaphors with sight as their source domain, antonymic gradable features like *bright* lead to higher accessibility than non-antonymic features like *red*.

Strik Lievers (2015) and Winter (2016) both discuss affectivity and lexical distribution, among other lexical factors which may influence directional preferences. They both relate metaphorical usage to affectivity, and propose that certain modalities, particularly taste and smell, are more affectively loaded than others. In other words, one of the points of using metaphors is affective evaluation, and since taste and smell are affectively loaded, they make better source domains. Like gradability, affectivity is a lexical factor: any given percept can be described with either high-affectivity or low-affectivity words. Winter (2016) presents a corpus study showing that a word's affectivity is a reliable predictor of its use in naturally occurring synesthetic metaphors, as are its frequency and iconicity.

Strik Lievers (2015) and Winter (2016) also both note that lexical coding in different modalities is not distributed evenly across lexical categories. For example, in English there are relatively few lexical adjectives associated with sound, but relatively many associated with sight (Strik Lievers & Winter 2018). Therefore, it's statistically more likely that we find a sight adjective modifying a sound noun, than a sound adjective modifying a sight noun. This point echoes the early suggestion by Ullmann (1957), that mappings from sight to sound are more frequent than ones from sound to sight because the domain of sight is lexically richer than that of sound. While this explanation accounts well for corpus findings, additional assumptions are required for it to account for experimental findings. A priori, there is no reason to expect general facts of lexical distribution to influence judgments about specific pairs of concepts, like the stimuli presented to participants in experiments. Naturally, lexical distribution is also a lexical factor.

Summarizing, several lexical factors have been shown to have an effect on directional preferences in synesthetic metaphors (Petersen et al. 2007, Winter 2016, Strik Lievers & Winter 2018). As of yet, there is no comparable evidence for an independent, direct effect of any perceptual factor. We may then entertain the possibility that there is no property of the sensory modalities, in and of themselves, which directly causes directional preferences. In other words, there might not be any criterion relevant for directional preferences by which the senses are ordered hierarchically. The logical consequence would be that what has so often been referred to as Ullmann's hierarchy of the senses may turn out to be descriptively adequate, but explanatorily inert. The so-called

hierarchy might be "explained away", partially or entirely, as an artifact of independent, idiosyncratic lexical factors.

Before we relegate Ullmann's hierarchy to an artifact, we might want to ask why perceptual factors have not been tested directly, let alone established empirically, in previous research. One reason for this may be that perceptual factors are more difficult to operationalize and manipulate than lexical factors, whether one is annotating naturally occurring synesthetic metaphors, or constructing experimental stimuli (see Winter 2019a). A second reason is that the effects of perceptual factors may be difficult to isolate, given the ubiquity of confounding lexical factors. Of course, all synesthetic metaphors are limited by the inventory of sensory words in the relevant language. If a lexical field is sparse, as is the case for smell in English for instance, this limitation can be quite severe (Majid & Burenhult 2014). But even a rich lexical field is limiting, because sensory words tend to have complex, idiosyncratic meanings. Finding a set of words which are associated with different senses but are otherwise comparable, i.e., not differentiated by lexical category, affectivity, gradability, morphological complexity, or other lexical factors, can be a formidable task.

To illustrate, consider intensity, one of the few candidates for a dimension that is straightforwardly analogous across the senses (Levinson & Majid 2014). As such, we might expect to find comparable lexical means for expressing high and low intensity in different sensory domains, yet we don't. In English, *dim*, *quiet* and *bland* can mean low intensity of light, sound and flavor, respectively, but they are not truly comparable because their meanings are more complex than that. This can be seen from their antonymy relations. *dim*, in addition to being an antonym of *bright*, is also an antonym of *clear*, which is completely orthogonal to intensity. The two obvious antonyms of *quiet* are *loud* and *noisy*, which both mean high intensity, but the latter also means something like erratic or disturbing. And the obvious antonym of *bland* is *tasty*, which means positive evaluation rather than high intensity.

The preceding paragraph focuses on adjectives, because adjective-noun phrases, such as *soft brightness*, are the most frequent and best-studied kind of synesthetic metaphor. But the difficulties in controlling for lexical factors in sensory words are not limited to adjectives. An exception to the ubiquitous focus on adjective-noun phrases in the literature is Shen & Gadir's (2009) experimental study of the Hebrew genitive construction *X shel Y*, which included concrete nouns with a salient

sensory feature (e.g., *sukar shel bosem* ‘sugar of perfume’), as well as abstract nouns derived from adjectives (e.g., *melixut shel digdugiut* ‘saltiness of ticklishness’). Despite not using adjectives, there are conspicuous lexical semantic differences in the materials: contrast the highly affective *siraxon* ‘stench’ with the low affectivity *taam* ‘flavor’, and the antonymic *kshixut* ‘rigidity’ with the non-antonymic *tsehivut* ‘yellowness’.

In light of the above, my goal here is to experimentally investigate directional preferences in synesthetic metaphors, while controlling for lexical factors which previous studies have not accounted for, and which have thus potentially warped the empirical picture. To my knowledge, this is the first attempt to take on this methodological challenge, and as a result, the first investigation of a potential direct effect of perceptual factors on directional preferences.

The results of the experiment reported below show that some directional preferences do surface when lexical factors are controlled for. It thus provides unprecedented evidence that perceptual factors play a role in determining directional preferences in synesthetic metaphors. However, the directional preferences found here do not add up to an overarching preference for mappings either upward or downward on Ullmann’s hierarchy of the senses.

3.2 Experimental study

This study was designed to test whether directional preferences in synesthetic metaphors arise in the absence of lexical factors. To that aim, I use synesthetic metaphors in a verbal analogy construction, wherein the target and source domains are each evoked by a copulative perception verb (henceforth CPV): *look*, *sound*, *smell*, *taste*, or *feel*. (52) lists naturally occurring examples of synesthetic metaphors in verbal analogies, retrieved from the enTenTen15 corpus on www.sketchengine.eu (Kilgarriff et al. 2014).

(52) a. …the painting looks like my music sounds.

b. Debussy can sound like Monet looks.

c. …a concoction that tastes like roses smell.

d. My insides felt like my garden looked.

e. This song sounds exactly like watching snow feels.

Using the verbal analogy construction to test directional preferences in synesthetic metaphors crucially relies on the assumption that this construction actually involves metaphorical mapping. This is a nontrivial assumption, which ties into the long-running debate on the relationship between metaphors and comparisons in general (e.g., Glucksberg & Keysar 1990, Chiappe & Kennedy 2001, Croft & Cruse 2004, Bowdle & Gentner 2005, Glucksberg & Haught 2006). I assume, following Steen et al. (2010), Wolff & Gentner (2011) and Gil & Shen (2021), that there is such a thing as a metaphorical comparison, in the sense that it involves unidirectional mapping of properties or inferences from one domain to another (though see Steen et al. 2010, pp. 92-96, for challenges in delineating domains in these cases). Pertinent evidence in support of this position is that comparisons between concepts in different domains exhibit directional preferences parallel to other metaphors, such as a preference for abstract and concrete concepts in subject and complement positions, respectively, rather than vice versa (Ortony 1979, Shen 1997, Porat & Shen 2017).

By using CPVs to explore synesthetic metaphors, I control for a number of lexical factors which have potentially muddled the results of previous studies, where sensory domains were evoked using nouns and adjectives. First and foremost, using CPVs makes it possible to control for lexical semantic factors such as affectivity and gradability. This is because (attributary uses of) CPVs have directly comparable and very lean semantic contributions: on its attributary use, a CPV restricts the interpretation of its complement to a subset of dimensions associated with a single sensory modality (see section 1.4). Importantly, what determines the utterance's affectivity, gradability, etc., is the complement, not the verb. To illustrate, contrast the highly affective and gradable (53a) with the low affect and non-gradability of (53b).

(53) a. The wines smell wonderful. (= The wines are wonderful with respect to their smell)
b. The wines smell identical. (= The wines are identical with respect to their smell)

If the CPV's complement does not have the dimensions specified by the verb, an attributary use is literally impossible, often resulting in an infelicitous utterance, e.g. *#smell purple*. However, contextual and grammatical cues can encourage or even impose a nonliteral reading on an utterance (Porat & Shen 2015). In the case of attributary uses of CPVs, a nonliteral interpretation may involve metaphorical mapping from the domain of the complement to the domain specified by the verb. If the complement evokes a particular sensory domain, the result may be a synesthetic

metaphor. A naturally occurring example of this is (54), where the phrase *For lack of a better word* and the double quotation marks around the complement serve to impose a nonliteral reading, which is then achieved via a mapping from sight to smell.

(54) For lack of a better word, it smells "purple".

Often, however, cases in which a CPV's complement evokes a different domain than the verb itself are naturally interpreted as evidential rather than attributary uses (Petersen & Gamerschlag 2014; see also sections 1.3-1.4). In addition, if a CPV's complement is a lexical predicate associated with a particular sensory domain, e.g., *purple*, we again run into issues of lexical coding and lexical semantics.

To control for lexical factors, and to guarantee an attributary use rather than an evidential use, I use the verbal analogy construction in (54a). In essence, the construction expresses that some implicit description which applies to stimulus *b* as perceived via modality *y* (the source domain), when mapped to modality *x* (the target domain), also applies to stimulus *a*. For example, (54b) expresses that an auditory description of the speaker's music, when mapped to the visual dimension, applies to the painting.

(55) a. NOUN-PHRASE$_a$ COPULATIVE-VERBS$_x$ like NOUN-PHRASE$_b$ COPULATIVE-VERBS$_y$

b. The painting looks like my music sounds.

An additional advantage of using CPVs is circumventing the issue of differential lexical distribution. This is because English CPVs comprise a closed set of lexemes, which stand in a one-to-one relation to the five Aristotelean senses. That is, each of the five sensory domains can be evoked using exactly one CPV, making them all equally encoded. As such, none of the domains is more or less likely to require the "recruitment" of descriptors from another domain, or to be recruited in the description of another domain.

Furthermore, using CPVs justifies some of the assumptions criticized by Winter (2016, 2019a) in earlier studies. Since there are exactly 5 English CPVs and each evokes a single sensory modality, the question of how to delineate the senses is resolved straightforwardly: I assume the 5 senses for which there are CPVs.[27] Likewise, each CPV can be assumed to be categorically associated with

[27] This coarse-grained categorization makes it prima facie impossible to assess the influence of any individual perceptual factor. The present study only tests for a potential effect of perceptual factors in

the modality it evokes, and not associated with any other modality. Lastly, synesthetic metaphors in the verbal analogy construction can be assumed to be categorically novel and not conventional, given their infrequency in natural usage (see section 3.3).

3.2.1 Materials

The experimental materials consisted of 80 short passages, each containing a synesthetic metaphor in the verbal analogy construction. Each passage consisted of (i) an explicit value judgment, i.e. *I (don't) like how this noun verbs*; followed by (ii) the phrase *In a way*; and finally (iii) a verbal analogy containing two different inanimate nouns, two different CPVs, and a modality-general adjective. The template for the passages is given in (55a), with an example in (55b).

(56) a. I (don't) like how this NOUN_a VERBS_x. In a way, this NOUN_a VERBS_x like a(n) ADJECTIVE NOUN_b VERBS_y.

b. I like how this coat feels. In a way, this coat feels like an expensive soup tastes.

The explicit value judgment and the abstract adjective were included as contextual cues for the interpretation of the verbal analogy, following a pilot experiment in which participants found bare verbal analogies difficult to interpret. The explicit value judgment also served to override differences in affective connotations between the CPVs. The phrase *In a way* was included to encourage a nonliteral interpretation of the verbal analogy.

Each synesthetic metaphor contained two different verbs, two different nouns, and one adjective. In total, 5 verbs, 40 nouns and 12 adjectives were used in the experiment. The verbs were the five English CPVs: *look*, *sound*, *smell*, *taste* and *feel*. For each verb, 8 inanimate, concrete nouns were chosen from among the 50 most frequent subjects of that verb occurring as a CPV (i.e., tagged as a verb, preceded by a word tagged as a noun, and followed by either a word tagged as an adjective or the word *like*), in the Sketch Engine corpus enTenTen15 (Kilgarriff et al. 2014). Each noun was only used as the subject of a single verb. For example, *car* was only used as a subject of *sound*, despite also being a frequent subject of *look* and *smell*.

The adjectives were chosen to represent 5 modality-general dimensions, 3 of which correspond to Osgood et al.'s (1957) affective components: *good*/*bad* for valence, *interesting*/*boring* for arousal,

aggregate. Even then, it potentially occludes the effects of certain perceptual factors, e.g., ones that differentiate the perception of texture and temperature, or the perception of color and shape.

and *strong*/*weak* for dominance. Half of the occurrences of *strong*/*weak* were substituted with *huge*/*tiny* to allow phrases which were otherwise difficult to interpret, e.g., *strong painting*, *weak house*. The remaining 2 dimensions were abstract: *familiar*/*strange* for familiarity, and *expensive*/*cheap* for price.

Nouns, verbs and adjectives were combined pseudo-randomly to create 40 synesthetic metaphors, such that each noun appeared in two metaphors with two different, non-antonym adjectives. From these 40 metaphors, another 40 metaphors were generated by flipping the order of the nouns and verbs. For example, *This coat feels like an expensive soup tastes* was flipped to create *This soup tastes like an expensive coat feels*. The 80 metaphors were embedded in the template in (55a) above, with the adjective's polarity determining whether the value judgment was positive or negative. The full list of metaphors is available online at https://osf.io/2hmcb/.

3.2.2 Procedure

48 monolingual English speakers were recruited over the Prolific crowd-sourcing platform. One participant showed zero variance in their responses, so their responses are excluded below, leaving 47 participants.

The experiment was created on www.qualtrics.com. Four lists were created, each consisting of 20 passages, with each of the 40 nouns appearing once per list, and the number of positive and negative value judgments counterbalanced between lists. The order of the passages in each list was randomized, and participants were randomly assigned to one of the four lists.

Participants were told they would see figurative sentences expressing opinions about things and comparing them to other, possibly very different things. They were instructed to rate how natural or unnatural each sentence was. A natural sentence was defined as "one that makes sense, that you would not be surprised to hear in conversation", an unnatural sentence as "one that doesn't make sense, and sounds awkward or foreign", and an intermediate sentence as "one you could make sense of, perhaps with some difficulty, though you might not expect to hear it in conversation". Participants rated the naturalness of each sentence on a 7-point scale, with 7 labelled "very natural" and 1 labelled "very unnatural".

Prior to the experiment proper, participants saw two practice questions designed to establish benchmarks of naturalness and unnaturalness. The first practice question included a conventional

metaphorical mapping, and the second included an anomalous metaphorical mapping. The two practice questions were followed by explanations tying them to the instructions and suggesting how their naturalness might be rated:

(57) I don't like how this person talks. In a way, this person talks like an excited puppy chases a thrown ball.

The writer of the sentence above expresses a negative opinion about the way a person talks, and then compares that to the way a puppy behaves.

Some people think this sentence makes a fair amount of sense (perhaps it means that the person talks in a way that is overexcited, childish, or lacks focus). Because of this, they may rate this sentence as quite natural.

(58) I like how this device works. In a way, this device works like a hungry shark searches for prey.

The writer of the sentence above expresses a positive opinion about the way a device works, and then compares that to the way a shark behaves.

Some people think this sentence doesn't make a lot of sense, or that it's hard to make sense of. Because of this, they may rate this sentence as rather unnatural.

3.2.3 Results

Figure 7 presents the overall distribution of the naturalness ratings. The distribution is centered below the middle of the naturalness scale (M = 3.1, SD = 1.63, median = 3, mode = 2), indicating that the participants generally found that the synesthetic metaphors didn't make much sense or were difficult to make sense of. This result is not surprising, given that the experimental materials were novel metaphorical mappings presented with little supporting context.

Figure 8 presents the means and interquartile ranges of the naturalness ratings grouped by sense combination and mapping direction. Of the 10 sense combinations, naturalness was by far highest in smell+taste (M = 4.48, SD = 1.66), which was also the only combination for which the mean as well as the median (= 5) were above the middle of the naturalness scale. For all 9 other sense combinations, both the mean and the median were below the middle of the scale. Further setting smell+taste apart from the other combinations, the mean difference between this pair and the next

highest combination, smell+feel (M = 3.44, SD = 1.71), was greater than the difference between the second highest and the very lowest-rated combination, look+sound (M = 2.58, SD = 1.63).

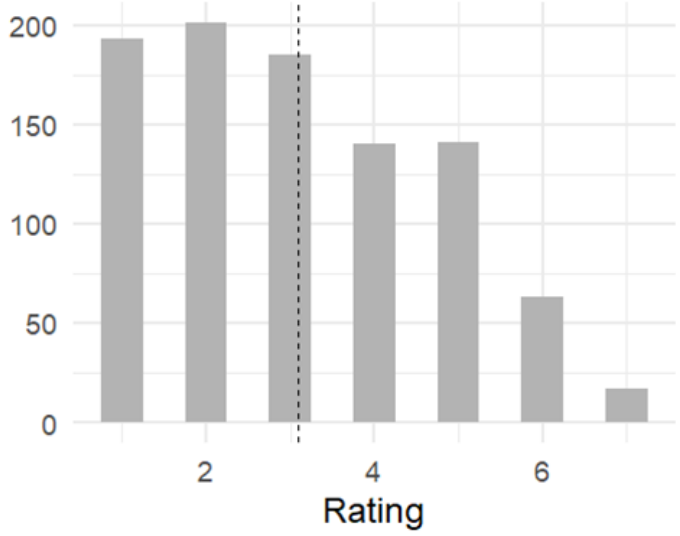


Figure 7. Overall distribution of naturalness ratings (on a scale of 1 through 7). The y-axis represents raw counts of responses grouped by rating. The dashed line indicates the grand mean.

Turning to mapping direction, naturalness across sense combinations was slightly higher in upward mappings (M = 3.15, SD = 1.69) than in downward mappings (M = 3.04, SD = 1.58). Within the 10 sense combinations, mean naturalness was higher in upward mappings than in downward mappings in 6 combinations, but lower in the remaining 4. The mean difference between upward and downward mappings was greatest in sound+feel, where upward mappings were preferred (M = 3.62, SD = 1.73; downward: M = 2.87, SD = 1.44). The next greatest difference was in look+sound, where the opposite direction was preferred (upward: M = 2.38, SD = 1.50; downward: M = 2.88, SD = 1.74).

The results were analyzed with a mixed-effects ordinal model. Analysis was conducted in the R software environment (using R version 3.6.3, R Development Core Team 2020), with the packages 'ordinal' (Christensen 2018) and 'tidyverse' version 1.3.0 (Wickham et al. 2019). Data were entered into a cumulative link model (i.e., ordinal regression model) with fixed effects for mapping direction (upward mapping and downward mapping), senses (each of the 10 possible 2-sense combinations), and value judgment (positive and negative), all of which were sum-coded. The

analysis also included an interaction term for direction × senses, and a random effect for participants. The scripts and the data are available online at https://osf.io/2hmcb/.

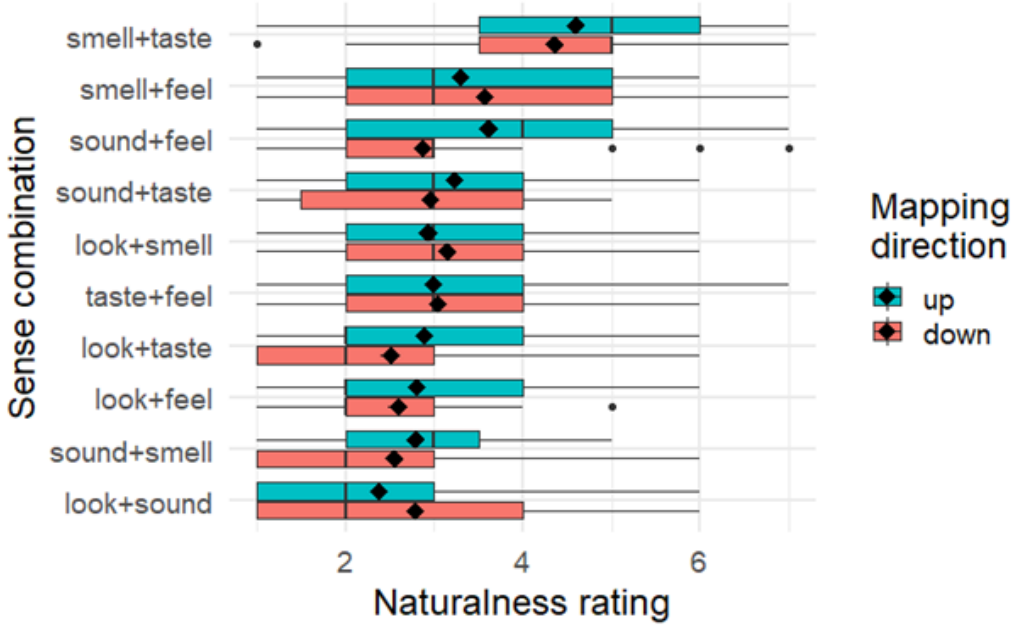


Figure 8. Distributions of naturalness ratings (on a scale of 1 through 7), grouped by sense combination and mapping direction. The boxes indicate the interquartile ranges, the vertical lines indicate medians, and the rhombuses indicate means. The whiskers extend to the interquartile range × 1.5 in each direction, or to the minimum/maximum values.

The effect of mapping direction on naturalness was minor, and not statistically significant (β = -0.09, *SE* = 0.06, p = .135). There was a significant effect of value judgment, with naturalness for negative judgments lower than the grand mean (β = -0.54, *SE* = 0.07, p < .001). There were also multiple significant effects of sense combination on naturalness: naturalness was considerably higher than the grand mean in smell+taste (β = 1.95, *SE* = 0.19, p < .001), and lower, to varying degrees, in look+sound (β = -1.12, *SE* = 0.19, p < .001), look+taste (β = -0.83, *SE* = 0.19, p < .001), and sound+smell (β = -0.59, *SE* = 0.18, p = .001).

The interaction between mapping direction and sense combination had a noticeable and statistically significant effect for two sense combinations: in look+sound, naturalness in downward mappings was higher than the mean (Estimate = 0.38, *SE* = 0.19, p = .038). Conversely, in sound+feel, naturalness in downward mappings was lower than the mean (Estimate = -0.35, *SE* = 0.18, p = .049).

3.3 Discussion

The results reported above suggest that localized directional preferences exist for synesthetic metaphors, even when lexical factors are controlled for. Specifically, there was a noticeable preference for downwards mappings in look+sound, and an opposite preference in sound+feel. These two opposite preferences align with two of Ullmann's (1945, 1957) early observations: first, that touch-to-sound mappings are the single most frequent type of synesthetic mapping; and second, that mappings between sight and sound are the single consistent exception to the general preference for mappings upwards on the hierarchy of the senses.

At the same time, the results provide no evidence that synesthetic mappings upwards on Ullmann's hierarchy of the senses are consistently preferred over downwards mappings. First, the overall effect of mapping direction was minor. Second, upwards mappings actually received lower mean naturalness ratings than downwards mappings in 4 out of 10 possible sense combinations. These results do not align with the findings of numerous earlier experimental studies (Shen & Cohen 1998, Shen & Eisenman 2008, Shen & Gil 2008, Shen & Gadir 2009). The design of the present experiment fundamentally differs from previous designs in two ways: (i) it evokes sensory domains using CPVs rather than adjectives or nouns; and (ii) it relates sensory domains using analogy rather than modification or predication. In the next section, I consider how these differences may be responsible for the contrast between present and past findings.

Next, the results indicate that mappings between certain senses, regardless of direction, are more natural than others. Particularly, mappings between smell and taste received considerably higher naturalness ratings than all other possible sense combinations, and were the only ones for which mean and median naturalness were higher than the midpoint of the scale.

A possible explanation for the gap between smell+taste and all the other sense combinations is that comparisons between smell and taste percepts are not actually metaphorical at all (see Fishman 2020, for comparable differences in ratings of literal and metaphorical comparisons). That is, the sensory domains of smell and taste may be so similar, or intersect to such an extent, that comparisons between them are naturally taken as intra-domain rather than cross-domain mappings. This idea is obliquely supported by the strong positive relationship between gustatory and olfactory measures of words (Lynott and Connell 2013, see also Winter 2016), as well as by neurocognitive evidence for integration between the gustatory and olfactory systems (Verhagen & Engelen 2006).

To test the reliability of the experimental results, I conducted a corpus study. I ran a search for the verbal analogy construction with CPVs in the Sketch Engine corpus enTenTen15 (Kilgarriff et al., 2014), using the following query:

> [lemma="look|sound|smell|taste|feel" & tag="V.*"] []{0,2} [lemma="like"] []{1,3} [lemma="look|sound|smell|taste|feel" & tag="V.*"]

The above query returns results which include the following five elements, in order: (i) one of the five lexemes *look*, *sound*, *smell*, *taste* and *feel*, tagged as a verb; (ii) a sequence of 0 to 2 words (for a potential modifier, e.g., *a bit*, *very much*, *exactly*, or a potential perceiver argument, e.g., *to me*); (iii) the word *like*; (iv) a sequence of 1 to 3 words (for the subject of the second verb and a potential auxiliary verb, e.g., *have*, *would*); and again (v) one of the five lexemes *look*, *sound*, *smell*, *taste* and *feel* tagged as a verb.

I extracted a random sample of 10,000 hits, which I then manually inspected to filter out false positives and duplicate hits, leaving 869 unique instances of the verbal analogy construction. Next, I hand-coded the sample, filtering out occurrences of the verbs as experiencer verbs rather than CPVs (i.e., with a perceiver rather than a stimulus as grammatical subject), as well as instances where one or both CPVs could be interpreted as evidential rather than attributary. This left 413 instances, presented in Table 11. Of these, the majority were instances containing the same CPV in both verbal positions, meaning they were literal intra-domain comparisons rather than synesthetic metaphors. Only 90 instances were actual cross-domain mappings, indicating that synesthetic metaphors in the verbal analogy construction are quite rare in natural usage. This finding aligns with the low mean naturalness ratings elicited in the experiment.

The small size of the corpus makes it impossible to draw statistically reliable conclusions. Perhaps the most conspicuous finding is the high number of mappings between smell and taste (n = 48), which account for over half of the cross-domain mappings. This fits with the substantial difference in naturalness ratings in the experiment, between smell+taste on the one hand, and all other sense combinations on the other hand. However, downward mappings from smell to taste were far more frequent in the corpus (43 of 48), whereas there was no clear preference for either direction in the experimental results.

Table 11. All instances of the verbal analogy construction with attributary uses. Upward and downward mappings are in the top right and bottom left halves, respectively. Literal comparisons are in the diagonal, in parentheses.

Target\Source	*look*	*sound*	*smell*	*taste*	*feel*	Total
look	(213)	6	0	0	4	223
sound	10	(43)	0	1	9	63
smell	0	0	(6)	5	1	12
taste	0	1	43	(22)	1	67
feel	4	4	0	1	(39)	48
Total	227	54	49	29	54	413

Upward mappings from touch to sound were somewhat more frequent than the opposite (9 of 13), as were downward mappings from look to sound (10 of 16). These findings are compatible with the directional preferences in the experiment, but the numbers are far too small to be reliable. The entirety of the corpus data is available online at https://osf.io/2hmcb/.

3.4 Conclusions

The experiment described above is, to my knowledge, the first experiment to probe directional preferences in synesthetic metaphors while controlling for lexical factors. As such, it constitutes the first attempt to directly test the effect of perceptual factors on directional preferences in synesthetic metaphors. This testing is made possible by focusing on an oft-overlooked set of perception verbs and using a novel construction: CPVs and the verbal analogy construction, respectively. This study thus circumvents a crucial methodological limitation of previous research into synesthetic metaphors, while also broadening the empirical scope of the phenomenon.

The results reveal directional preferences in verbal analogies with CPVs, namely in mappings between touch and sound, and between sight and sound. Although more localized than the overarching preference for upward mappings observed in earlier studies, these preferences do align with previous findings. Mappings from touch to sound, and from sight to sound, have consistently been found to be preferred over their opposites, and generally rank among the top possible

mappings in frequency, accessibility, and comprehensibility (Ullmann 1945, Williams 1976, Shen & Gil 2008, Shinohara & Nakayama 2011, Strik-Lievers 2015, Winter 2016). Importantly however, the present findings are the first that cannot be attributed to differences in lexical semantics or lexical coding. I would go further and venture that this is the first evidence for a direct effect of perceptual factors, i.e., properties of the sensory modalities themselves, on directional preferences in synesthetic metaphors. The experiment was not designed to explore which perceptual factors these may be, so I refrain from speculating on the matter. Nonetheless, these findings place new and important restrictions on any future theory of synesthetic metaphors, and, I believe, also point to exciting new avenues for future empirical research.

Perhaps the more striking finding arising from the present results is the lack of an overarching effect of mapping direction. In this, the present study diverges from decades of empirical research into synesthetic metaphors, comprising corpus and experimental studies in various languages, and consistently showing a preference for mappings "upwards" on Ullmann's hierarchy of the senses. How can we account for this divergence?

I propose that the preference for upward mappings observed in previous studies is due to one or more factors that are not in effect, or are somehow mitigated, in the present study. Given the goals of this study, some immediate suspects are lexical semantic factors, such as affectivity and gradability, along with differences in lexical coding. I have argued that these factors don't differentiate CPVs, and hence are rendered inert in the experiment reported here. At the same time, such factors have previously been shown to reliably predict the frequency and acceptability of synesthetic metaphors (Petersen et al. 2007, Winter 2016). It's not a huge leap to posit that the overarching preference for upward mappings is due to the accumulated effects of several such lexical factors, some of which we may not yet know about. If that is indeed the case, we might conclude that mapping direction with respect to Ullmann's hierarchy of the senses is an artifact, with no independent effect on synesthetic metaphors. Put another way, what appears to be a preference for upward mappings, may actually turn out to be a conflation of several independent lexical factors, which just so happens to (roughly) fit the ideas of classical philosophers about the senses.[28]

[28] As noted in section 3.1, some lexical factors may be traced back to perceptual factors, in that the properties of each sensory modality influence the makeup of the lexical inventory associated with that modality. These

It's also possible that certain factors relevant to directional preferences were unintentionally mitigated in the present study. Here I consider three such factors: the (im)possibility of metaphorical mapping in comparisons, the inherent directionality of the grammatical form, and the interpretability of the synesthetic metaphor.

As noted in section 3.2, I assume here that comparisons in general, and the verbal analogy construction in particular, may involve metaphorical mapping. I take the directional preferences revealed in the present study, which parallel two of the directional preferences most consistently observed with other synesthetic metaphors, as further evidence in support of this assumption. However, let us consider the alternative, that metaphorical mapping is fundamentally impossible in comparisons. Proponents of this view might argue that the reason the present study did not find additional directional preferences, e.g., an overarching preference for upward metaphors, is that comparison and metaphor are subject to influence by different perceptual factors. More specifically, it would seem that comparisons are influenced by a subset of the perceptual factors which influence metaphors: those that drive preferences for touch-to-sound and sight-to-sound, but not those that drive the general preference for upward metaphors. This then raises questions regarding which perceptual factors influence which figures of speech, and why they influence one but not the other.

The directionality of a grammatical form is the degree to which the form constrains the direction of metaphorical mapping (see Porat & Shen 2017, Gil & Shen 2021, Fishman & Shen in preparation). Adjectival modification (e.g., *soft brightness*) and nominal predication (e.g., *brightness is softness*) both exemplify high directionality, with strict mapping from adjective to noun and from predicate to subject, respectively. As such, preferences in mapping direction can be clearly detected using naturalness ratings about these constructions. Conversely, genitive constructions (e.g., *a softness of a brightness*) and comparisons in intransitive collective constructions (e.g., *softness and brightness are alike*) exemplify low directionality, with mapping direction virtually unconstrained. As such, preferences in mapping direction might be entirely

would be indirect effects of perceptual factors, mediated through a language's lexicon. Nothing in the present study precludes a hierarchy of the senses with such indirect effects on directional preferences. However, cross-linguistic evidence shows that coding of sensory words differs considerably across languages, suggesting that such a hierarchy would be language-specific rather than universal (Majid et al. 2018).

obfuscated in naturalness ratings about these constructions, though they can be revealed using other experimental tasks (Shen & Gadir 2009). The directionality of comparisons with a subject and a complement (e.g. *brightness is like softness*) is a matter of debate (e.g., Glucksberg & Keysar 1990, Chiappe & Kennedy 2001, Wolff & Gentner 2011), but plausibly falls somewhere between those two extremes. The verbal analogies used here are such comparisons, and therefore might be less inherently directional than other frequent forms of synesthetic metaphor, especially adjective-noun phrases. Hence, it's possible that there was a preference for upward mappings in the experiment reported here after all, but it went undetected due to the verbal analogy's relatively low directionality and the nature of the experimental task.

Another factor which may have stymied directional preferences in the experiment reported here is interpretability. A study by Fishman & Shen (in preparation) suggests that interpretability has an independent effect contributing to directional preferences. Fishman and Shen conducted an experiment testing preference between two grammatical forms of comparisons: an intransitive collective construction (*A and B are alike*; low directionality) and a construction with a subject and a complement (*A is like B*; higher directionality). They reasoned that speakers would choose the more directional form when they had a clearer preference for a particular mapping direction. They found a greater preference for the more directional form in interpretable metaphorical comparisons (e.g., *Salesmen are like bulldozers*) relative to anomalies, i.e., uninterpretable metaphorical comparisons (e.g., *Deserts are like bulldozers*). They conclude that interpretability, though not a necessary condition for directional preferences, plays a role independently of factors like concreteness and typicality. The present findings, namely the low observed frequency of synesthetic metaphors in the verbal analogy construction, and the overall low naturalness ratings elicited for the experimental materials, indicate that synesthetic metaphors in this construction are quite difficult to interpret. It may be that this difficulty stymies the preference for upward mappings relative to more frequent and more interpretable synesthetic metaphors, e.g., adjective-noun phrases. This is especially true if many of the latter rely on conventionalized meanings, as argued by Winter (2019b).

In spite of the present study's limitations with regard to directionality and interpretability, as discussed above, the experiment did find empirical evidence for some directional preferences. This is not to say that the issues of directionality and interpretability should be brushed off. On the

contrary, future research should address these issues directly, by investigating the interplay between these two factors and directional preferences, not only in synesthetic mapping but in metaphorical mapping more generally. I believe further exploration of these notions is crucial for advancing our understanding of metaphor.

4 General discussion

4.1 Summary

In this dissertation I have explored questions in the linguistic domains of perception, evidentiality, and metaphor. My research revolved around the class of copulative perception verbs (CPVs), which were defined as perception verbs taking a stimulus as their grammatical subject, and requiring a predicate or clausal complement (Viberg 1983, 2019). At the outset, CPVs were recognized as having two distinct meanings: an attributary meaning, wherein the verb attributes a property to a perceptual impression, and a parenthetical meaning, wherein the verb modifies a proposition, relating it to the speaker's evidence and/or their epistemic status (Rogers 1974, Lasersohn 1995, Gisborne 2010, Poortvliet 2018, Muñoz 2019, inter alia).

The first chapter of the dissertation was an empirical investigation of CPVs in Hebrew. I focused on three constructions that host Hebrew CPVs, and found correspondences between them and the aforementioned distinct meanings of CPVs. Hebrew CPVs stand out from CPVs in languages where they have been previously studied, in that they can take both adjectival and adverbial complements, in what I call the copulative and the verbal constructions, respectively. I conducted a large-scale corpus study and, using Distinctive Collexeme Analysis (Gries & Stefanowitsch 2004), I showed that these two constructions attracted different subjects and complements. Specifically, the copulative construction attracted abstract subjects (e.g., 'question') and abstract complements (e.g., 'logical'), attesting to a parenthetical meaning, whereas the verbal construction attracted perceivable subjects (e.g., 'person') and multidimensional subjective complements (e.g., 'good'), attesting to an attributary meaning.

In addition, I conducted a preference experiment to identify semantic differences between the copulative construction and the impersonal construction, where the verb takes a clausal complement and occurs either without a subject or with an expletive subject. The results showed that the copulative construction was preferred in contexts contradicting an epistemic meaning (e.g., 'I know that x is not P, but x looks P'), whereas the impersonal construction was preferred in contexts contradicting an evidential meaning (e.g., 'I've never seen x, but it looks like x is P'). These findings indicate that the copulative construction and the impersonal construction encode an evidential meaning and an epistemic meaning, respectively. Thus, each of the three studied constructions appears to have a distinct meaning.

In the second chapter I addressed the phenomenon of evidential uncertainty. Evidential uncertainty refers to an inference of speaker uncertainty, arising from the use of evidential expressions, that is, expressions which indicate what evidence a speaker has for a proposition (Pogue & Tanenhaus 2018, Degen et al. 2019). I introduced a novel hybrid account of evidential uncertainty, combining elements of two existing accounts: one which attributes the uncertainty inference to extralinguistic reasoning about evidence type and directness (von Fintel & Gillies 2010, Mandelkern 2019), and one which attributes it to Gricean reasoning (Degen et al. 2015). I also introduced a set of utterances for which the three accounts make different predictions, namely utterances where an evidential expression indicates the maximally direct evidence for a proposition. My case study was utterances in which the evidential expression was a CPV with a complement strongly associated with the same sensory modality as the CPV (e.g., *looks blue*).

To test the different predictions of the three accounts, I conducted two experiments in the paradigm developed by Degen et al. (2019). The first experiment, which placed participants in a listener's role and asked them to rate a speaker's certainty, revealed that evidential expressions consistently conveyed uncertainty relative to bare utterances without evidential expressions. Moreover, uncertainty was greater for evidential expressions indicating both indirect and maximally direct evidence. The second experiment placed participants in a speaker's role and asked them to choose between utterances with and without evidential expressions. Results here revealed that speakers were more likely to indicate maximally direct evidence in contexts which made that type of evidence unreliable.

I argued that the experimental findings support my hybrid account, which I formalized and implemented in a computational model within the Rational Speech Act framework (Frank & Goodman 2012). My model extended the basic framework with a representation of evidence strength, comprised of evidence directness and evidence reliability, which factored into the speaker's belief function and were included in the set of inferences outputted by the pragmatic listener function. With this added representation, the computational model was able to capture the critical qualitative findings of the two reported experiments.

In the third chapter I investigated synesthetic metaphors, which are metaphorical mappings between sensory domains, and specifically the phenomenon of directional preferences, whereby mappings from one domain to another are preferred over mappings in the opposite direction

(Ullmann 1945, 1957, Shen 1997, Strik Lievers 2015, inter alia). I classified the factors which have previously been proposed to drive directional preferences as either lexical or perceptual. Lexical factors are those that are dependent on particular word choices, and include gradability (Popova 2005, Petersen et al. 2007), affectivity (Winter 2016), and differential coding across lexical categories (Strik Lievers & Winter 2018). In contrast, perceptual factors are independent from particular word choices, and include degree of embodiment (Shen 1997, Shen & Eisenman 2008) and sensory association (Shibuya et al. 2007). I proposed that previous studies attempting to test the effects of perceptual factors have failed to adequately control for lexical factors, in large part due to idiosyncrasies in the meanings of sensory adjectives and nouns (cf. Winter 2019a).

I argued that using CPVs in an analogy construction (e.g., *The picture looks like my music sounds*) would allow us to control for lexical factors in a way that previous studies have been unable to, and thus empirically probe the effect of perceptual factors. I did this by conducting a naturalness rating experiment, which revealed directional preferences between touch and sound, as well as between sight and sound. However, the experiment revealed no evidence for any "global" directional preferences, as predicted by Ullmann's (1945, 1957) idea of a hierarchy of the senses. I additionally conducted a small-scale corpus study, providing converging evidence for the same directional preferences revealed in the experiment.

Taking a bird's-eye view of the thesis in its entirety, I have essentially outlined the full range of uses of CPVs. In the broadest terms, CPVs are a way of combining a given sensory modality with just about any property expressible in a language. Their most straightforward use appears to be describing a perceptual impression in general and rather vague terms, by combining the verb with a multidimensional property, as shown in chapter 1. This is because combining the verb with a unidimensional property already associated with the relevant sensory modality would be redundant, and would thus paradoxically convey uncertainty, as shown in chapter 2.

Combining a CPV with an abstract property, or with a property more strongly associated with some other sensory modality, is a typical way to mark source of information, aimed to present a proposition without committing too strongly to its truth. This is shown in chapters 1 and 2. Lastly, a CPV can also be combined with a unidimensional property associated with some other sensory modality, as an instance of figurative language. Such instances are rare, and their interpretability depends on the specific properties and sensory modalities involved, as shown in chapter 3.

4.2 Contributions

I believe that this dissertation has important implications for future studies in each of the domains touched upon here. With regard to our understanding of CPVs per se, this work provides empirical evidence for a formal correspondent to the semantic distinction between attributary and parenthetical meanings. This naturally leads to the hypothesis that a similar correspondence could be found in other languages, particularly ones with an alternation between copulative and verbal constructions.

With regard to the Hebrew verbal construction specifically, I proposed to analyze it as interacting directly with the dimensions of its complement (cf. Alrenga 2010, Sassoon 2013), to capture how the complement slot is virtually restricted to multidimensional predicates. I also tentatively proposed extending this analysis to attributary meanings in languages which don't have the formal alternation between copulative and verbal constructions, such as English. This analysis would be more restrictive than existing accounts of the attributary meaning (e.g., Petersen & Gamerschlag 2014, Muñoz 2019), in at least two ways. First, it would disallow unidimensional complements attributable to the relevant sensory modality (e.g., *look oblong*, *sound loud*, *feel smooth*), since the CPV's contribution would be redundant in such cases. Second, it would disallow "goal-oriented" complements, since they do not seem to be multidimensional (cf. *#good for making people queasy, with respect to smell*).

Next, I presented experimental evidence that the Hebrew copulative and impersonal constructions encode subtly different parenthetical meanings. The distinction seems to fall along the lines of evidential and epistemic modal meanings, casting doubt on the idea that the class of evidentials and the class of epistemic modals are one and the same (cf. Matthewson 2012). In addition, my findings again naturally lead to a hypothesis about other languages. The weaker version of this hypothesis is that copulative and impersonal constructions would be preferred in different contexts, attesting to a difference in their encoded meanings. The stronger version of the hypothesis is that copulative constructions across languages would tend to encode evidential rather than epistemic meanings, and vice versa for impersonal constructions.

Moving on from the evidential meaning of CPVs to evidential expressions more generally, my experimental findings pose challenges to existing accounts of evidential uncertainty (e.g., von Fintel &Gillies 2010, Degen et al. 2015). The finding that evidential expressions always convey

uncertainty on the speaker's part, even when indicating maximally direct evidence, weakens purely extralinguistic accounts. At the same time, the finding that evidence directness influences the degree of uncertainty weakens purely Gricean accounts. To capture the emerging empirical picture, I believe that our best choice is to posit a hybrid account, involving both extralinguistic reasoning and Gricean reasoning.

I also presented experimental evidence that the choice to use an evidential expression depends on both evidence directness and evidence reliability. Specifically, choosing to indicate maximally direct evidence is more likely when context makes the evidence poor. These findings corroborate Grice's doubt-or-denial condition (Grice & White 1961), as well as support a complex conception of evidence strength. I demonstrated how a formal representation of evidence strength along these lines can be incorporated into a model within the Rational Speech Act framework (Frank & Goodman 2012), and spelled out a set of assumptions which allow such a model to derive both evidential uncertainty and the doubt-or-denial condition.

In the final experimental study reported here, concerning directional preferences in synesthetic metaphors, I found directional preferences in mappings between touch and sound, and between sight and sound. These findings support multiple previous studies showing that mappings from touch to sound and from sight to sound are among the most frequent and best rated of all possible synesthetic mappings (Ullmann 1945, Williams 1976, Shen & Gil 2008, Shinohara & Nakayama 2011, inter alia). But, considering the methodological limitations of previous experimental and corpus studies, I take the present findings to be the first empirical evidence for a direct effect of perceptual factors on directional preferences.

It is perhaps more surprising that directional preferences were not found in other mappings, and in particular, that there was no overarching preference for mappings on either direction of Ullmann's purported hierarchy of the senses. These results lend further support to critiques of Ullmann's hierarchy and the rich literature following it (Winter 2016, 2019a, 2019b). Taken to their logical extreme, these critiques suggest that Ullmann's hierarchy of the senses is at best an imperfect grouping of several independent factors, and at worst a red herring, with no explanatory power for directional preferences in synesthetic metaphors.

I believe that this thesis also makes some small contribution to the way that linguistic inquiry is undertaken. It demonstrates the value of combining methodological tools and theoretical

approaches, on the one hand, and on the other hand, reveals the pitfalls of clinging to a single narrow view. Different empirical research questions can each be approached with a different methodology best-suited to it, whether corpus-based, experimental, or computational, and a single research question can be addressed far more convincingly with converging evidence from multiple methodologies. Similarly, ideas and insights from different theoretical frameworks can be integrated into novel accounts. Conversely, ignoring other frameworks can lead to missed observations and even unexpected flaws in arguments and empirical studies.

4.3 Future work

As with any academic work, this dissertation does not provide the final word on its research topics, and many of its contributions call for elaboration, corroboration, or both. First and foremost, CPVs warrant more attention, especially empirical investigation, and particularly in non-European languages. I hope that the present study has demonstrated something of their versatility as linguistic expressions, their complexity as an object of study, and the fruitfulness of leveraging their unique features to address other research topics.

As noted above, the findings of this dissertation raise a number of hypotheses regarding CPVs across languages, which could constitute potential avenues for future research. The first is that the correspondence observed in Hebrew, between the copulative and verbal constructions on one hand, and the parenthetical and attributary meanings on the other hand, would be found in other languages where CPVs can take both adjectival and adverbial complements. We already know of one other language, Russian, where this hypothesis could be empirically tested, using the exact same techniques for corpus analysis employed here. Given the relative dearth of research into CPVs in non-European languages, there may be many more languages where this hypothesis would be applicable.

The second hypothesis, arising from the finding that the Hebrew verbal construction requires a multidimensional complement, is that the attributary meaning in other languages would be similarly restricted, perhaps even in languages without a specialized verbal construction, such as English. As outlined above, this hypothesis could be tested either with unidimensional complements attributable to the relevant sensory modality, or with “goal-oriented” complements. In a language like English, determining whether an attributary meaning is possible in such cases may be difficult due to the virtually unrestricted availability of the parenthetical meaning, but this

difficulty could potentially be overcome with careful experimental design (cf. Hansen & Markman 2005).

Next is the hypothesis that, just as copulative and impersonal constructions exhibit semantic differences in Hebrew, they would exhibit the same or similar differences cross-linguistically. The preference experiment conducted here could be reproduced in other languages. The experimental paradigm could even be expanded to include the so-called "copy raising" construction, which takes a subject just like the copulative construction, and a clausal complement just like the impersonal constructions. Doing so would allow us to test whether the subject and the complement each have their own contribution to the evidential and epistemic meaning components, or if the latter are determined by the construction as a whole.

Beyond testing the hypotheses raised here, future research could also attempt to shore up, or alternatively, to challenge the weak points of each of the experimental studies I reported. The preference experiment exploring the semantics of copulative and impersonal constructions was limited to only two CPVs, and accordingly, to only two types of evidence. Moreover, I tested only auditory evidence for the CPV *nishma* 'sound', yet the results, as well as previous literature, suggest that reported evidence is equally relevant (e.g., Viberg 2019). Other CPVs may allow other types of non-perceptual evidence. Future research may reveal that the impersonal construction does encode evidential meaning after all, albeit less specific or restricted than that encoded by the copulative construction.

The present experiments on evidential uncertainty have their limitations as well. The interpretation experiment explored only a single CPV, and a rather narrow range of evidence directness values. The production experiment presented participants with a forced choice between an utterance with an evidential expression and a bare utterance without one, even in contexts where both options were inappropriate. These limitations precluded a quantitative evaluation of the computational model. In future research I aim to conduct new and improved variants of both experiments, so as to reinforce the findings and also allow quantitative comparisons with, as well as fine-tuning of, the computational model's predictions.

The naturalness rating experiment on synesthetic metaphors addressed the shortcomings of many previous studies by controlling for lexical factors. However, it differs from earlier experimental studies in another major way, namely in relating sensory domains using analogy rather than

modification or predication. As such, the differences between past and present findings could be attributed to the presence of analogy, instead of to the absence of lexical factors. I see no obvious way to replicate the present experiment without an analogy construction. Therefore, I believe that the best way to address this potential criticism would be to replicate earlier studies, which is to say, studies which did not control for the relevant lexical factors, using analogies in the place of adjective-noun phrases or genitive constructions.

Finally, future research could build on the contributions of this dissertation, and address follow-up questions not addressed here. In my view, the biggest gap in our understanding of CPVs, at this time, is occurrences with comparative complements. We know that they can have both attributary and parenthetical meanings, but there is almost no data on these constructions. Future work on CPVs with comparative complements could discuss each of the domains explored here, starting from the very basics of their use, to the way they convey uncertainty, to (additional) cases of metaphorical (or analogical) mapping. How are CPVs with comparative constructions similar to or different from copulative and verbal constructions? When do speakers choose to use a comparative complement rather than a lexical predicate? Is the standard of comparison (e.g., the complement of *like*) prone to be generic or specific, or does this depend on the intended meaning?

Other future studies could address more specific follow-up questions. For example, we now know that speakers use evidential expressions when they have one type of available evidence, which is either indirect or unreliable. How do they choose which evidential expression to use if they have several, possibly conflicting types of available evidence? Similarly, we now know that mappings from touch to sound, and from sight to sound, are preferred over mappings in the opposite direction. Which perceptual factors drive these preferences? Are they related to the source domain or the target domain? These questions and many others currently remain unanswered, but I hope that we are at least one step closer to answering them.

References

Aikhenvald, Alexandra Y. (2004). *Evidentiality*. Oxford, UK: Oxford University Press.

Aikhenvald, Alexandra Y. (2014). The grammar of knowledge: A cross-linguistic view of evidentials and the expression of information source. In Alexandra Y. Aikhenvald & Robert M. W. Dixon (eds.), *The grammar of knowledge: A cross-linguistic typology*: 1-51. Oxford, UK: Oxford University Press.

Aikhenvald, Alexandra Y. & Anne Storch. (2013). Linguistic expression of perception and cognition: A typological glimpse. In Alexandra Y. Aikhenvald & Anne Storch (eds.), *Perception and cognition in language and culture*: 1-46. Leiden, NL: Brill.

Albelda Marco, Marta & Marlies Jansegers. (2019). From visual perception to evidentiality: A functional empirical approach to se ve que in Spanish. *Lingua* 220, 76-97.

Alrenga, Peter. (2010). Comparisons of similarity and difference. In Patricia Cabredo Hofherr & Ora Matushansky (eds.), *Adjectives: Formal analyses in syntax and semantics*: 155-186. Amsterdam, NL: John Benjamins Publishing Company.

Anand, Pranav & Natasha Korotkova. (2018). Acquaintance content and obviation. In Uli Sauerland & Stephanie Solt (eds.), *Proceedings of Sinn und Bedeutung 22*, Volume 1, ZAS Papers in Linguistics 60, 55-72.

Asudeh, Ash. (2002). Richard III. In Mary Andronis, Erin Debenport, Anne Pycha & Keiko Yoshimura (eds.), *Papers from the 38th regional meeting of the Chicago Linguistic Society*, 31-46. Chicago, IL: Chicago Linguistic Society.

Asudeh, Ash. (2004). *Resumption as resource management*. Stanford, CA: Stanford University dissertation.

Asudeh, Ash & Ida Toivonen. (2012). Copy raising and perception. *Natural Language & Linguistic Theory* 30(2), 321-380.

Avineri, Bar. (2021). Alternating smell in Modern Hebrew. In Łukasz Jędrzejowski & Przemysław Staniewski (eds.), *The linguistics of olfaction*: 305-342. Amsterdam, NL: John Benjamins Publishing Company.

Bowdle, Brian F. & Dedre Gentner (2005). The career of metaphor. *Psychological Review* 112(1), 193-216.

Chiappe, Dan L. & John M. Kennedy (2001). Literal bases for metaphor and simile. *Metaphor and Symbol* 16(3-4), 249–276.

Christensen, Rune Haubo B. (2018). *Ordinal – Regression models for ordinal data*. R package version 2018, 8–25. http://www.cran.r-project.org/package=ordinal/.

Croft, William & D. Alan Cruse (2004). *Cognitive linguistics*. Cambridge, UK: Cambridge University Press.

Day, Sean. (1996). Synaesthesia and synaesthetic metaphors. *Psyche* 2(32), 1-16.

Degen, Judith & Noah D. Goodman. (2014). Lost your marbles? The puzzle of dependent measures in experimental pragmatics. *Proceedings of the Annual Meeting of the Cognitive Science Society* 36, 397-402.

Degen, Judith, Justine T. Kao, Gregory Scontras & Noah D. Goodman. (2015). A cost and information-theoretic account of epistemic "must". Poster presented at *CUNY 2015*, March 19-21, Los Angeles, CA.

Degen, Judith, Andreas Trotzke, Gregory Scontras, Eva Wittenberg & Noah D. Goodman. (2019). Definitely, maybe: A new experimental paradigm for investigating the pragmatics of evidential devices across languages. *Journal of Pragmatics* 140, 33-48.

Dombi, Erzsébet P. (1974). Synaesthesia and poetry. *Poetics* 11, 23-44.

Evans, Nicholas & David Wilkins. (2000). In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language* 76(3), 546-592.

Faller, Martina T. (2019). The discourse commitments of illocutionary reportatives. *Semantics and Pragmatics* 12(8), 1-46.

Faller, Martina T. (2020). A possible worlds semantics for Cuzco Quechua evidentials. In Chungmin Lee & Jinho Park (eds.), *Evidentials and Modals*: 41-68. Leiden, NL: Brill.

von Fintel, Kai & Anthony S. Gillies. (2010). Must… stay… strong! *Natural Language Semantics* 18, 351-383.

Fernández Jaén, Jose. (2015). The pseudo-copulative verbs verse and sentirse: Diachronic and synchronic aspects. In Elisa Barrajón López, José Luis Cifuentes Honrubia & Susana Rodríguez Rosique (eds.), *Verb classes and aspect*: 218-237. Amsterdam, NL: John Benjamins Publishing Company.

Fishman, Alon. (2020). English similarity predicates construe particular dimensions of similarity. *Cognitive Linguistics* 31(3), 453-484.

Fishman, Alon & Yeshayahu Shen. (In preparation). *The (bidirectional) relation between grammatical asymmetry and metaphoricity*. Unpublished manuscript, Tel Aviv University.

Frank, Michael C. & Noah D. Goodman. (2012). Predicting pragmatic reasoning in language games. *Science* 336: 998.

Gegenfurter, Karl R., Marina Bloj & Matteo Toscani. (2015). The many colours of 'the dress'. *Current Biology* 25(13), R543-R544.

Gil, David & Yeshayahu Shen. (2021). Metaphors: The evolutionary journey from bidirectionality to unidirectionality. *Philosophical Transactions of the Royal Society B* 376(1824).

Gisborne, Nikolas. (2010). *The event structure of perception verbs*. Oxford, UK: Oxford University Press on Demand.

Glucksberg, Sam & Boaz Keysar. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review* 97(1), 3-18.

Glucksberg, Sam & Catherine Haught. (2006). On the relation between metaphor and simile: When comparison fails. *Mind & Language* 21(3), 360-378.

Goodman, Noah D. & Andreas Stuhlmüller. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, *5*(1): 173-184.

Goodman, Noah D. & Michael C. Frank. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences* 20, 818–829.

Grice, H. Paul & Alan R. White. (1961). The causal theory of perception. *Proceedings of the Aristotelian Society, Supplementary Volumes* 35, 121-152.

Grice, H. Paul. (1989). *Studies in the Ways of Words*. Cambridge, MA: Harvard University Press.

Gries, Stefan Th. & Anatol Stefanowitsch. (2004). Extending collostructional analysis: A corpus-based perspective on "alternations." *International Journal of Corpus Linguistics* 9(1), 97-130.

Gries, Stefan Th. (2014). Coll.analysis 3.5. A script for R to compute perform collostructional analyses. https://www.stgries.info/teaching/groningen/index.html.

Gisborne, Nikolas. (2010). *The event structure of perception verbs*. Oxford, UK: Oxford University Press.

Hansen, Mikkel B. & Ellen M. Markman. (2005). Appearance questions can be misleading: A discourse-based account of the appearance-reality problem. *Cognitive Psychology* 50(3), 233-263.

Heycock, Caroline B. (1994). *Layers of predication: The non-lexical syntax of clauses*. New York, NY: Garland Publishing.

Higginbotham, James. (1983). The logic of perceptual reports: An extensional alternative to situation semantics. *The Journal of Philosophy* 80(2), 100-127.

Huumo, Tuomas. (2010). Is perception a directional relationship? On directionality and its motivation in Finnish expressions of sensory perception. *Linguistics* 48(1), 49-97.

Ibarretxe-Antuñano, Iraide. (1999). *Polysemy and metaphor in perception verbs: A cross-linguistic study*. Edinburgh, UK: University of Edinburgh dissertation.

Jackendoff, Ray. (2007). *Language, consciousness, culture: Essays on mental structure*. Cambridge, MA: The MIT Press.

Kaiser, Elsi. (2018). Effects of sensory modality on the interpretation of subjective adjectives: Comparing sight, smell and taste. *Proceedings of the Berkeley Linguistics Society* 44, 99-114.

Kaiser, Elsi & Deniz Rudin. (2020). When faultless disagreement is not so faultless: What widely-held opinions can tell us about subjective adjectives. *Proceedings of the Linguistic Society of America* 5(1), 698-707.

Kao, Justine T., Leon Bergen & Noah D. Goodman. (2014a). Formalizing the pragmatics of metaphor understanding. *Proceedings of the Annual Meeting of the Cognitive Science Society* 36, 719-724.

Kao, Justine T., Jean Y. Wu, Leon Bergen & Noah D. Goodman. (2014b). Nonliteral understanding of number words. *Proceedings of the National Academy of Sciences* 111(33), 12002-12007.

Karttunen, Lauri. (1972). Possible and must. In John P. Kimball (ed.), *Syntax and Semantics, Volume 1*: 1-20. New York, NY: Academic Press.

Kennedy, Chris. (2013). Two sources of subjectivity: Qualitative assessment and dimensional uncertainty. *Inquiry* 56(2-3), 258-277.

Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý & Vít Suchomel. (2014). The Sketch Engine: Ten years on. *Lexicography* 1, 7-36.

Kölbel, Max. (2004). Faultless disagreement. *Proceedings of the Aristotelian society* 104(1), 53-73.

Korotkova, Natasha. (2020). Evidential meaning and (not-)at-issueness. *Semantics and Pragmatics* 13, 1-26.

Kratzer, Angelika. (1991). Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research* 6: 639–650. Berlin, DE: Walter de Gruyter.

Kurumada, Chigusa, Meredith Brown, Sarah Bibyk, Daniel F. Pontillo & Michael K. Tanenhaus. (2014). Is it or isn't it: Listeners make rapid use of prosody to infer speaker meanings. *Cognition* 133, 335-342.

Lakoff, George & Mark Johnson. (1980). The metaphorical structure of the human conceptual system. *Cognitive Science* 4(2), 195-208.

Lakoff, George & Mark Johnson. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York, NY: Basic Books.

Landau, Idan. (2011). Predication vs. aboutness in copy raising. *Natural Language & Linguistic Theory* 29(3), 779-813.

Lasersohn, Peter. (1995) Sounds like like. *Linguistic Analysis* 25, 70-77.

Lasersohn, Peter. (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy* 28(6), 643-686.

Lassiter, Daniel & Noah D. Goodman. (2013). Context, scale structure, and statistics in the interpretation of positive-form adjectives. *Proceedings of Semantics and Linguistic Theory* 23, 587–610.

Lassiter, Daniel. (2016). Must, knowledge, and (in)directness. *Natural Language Semantics* 24, 117-163.

Late Night with Seth Meyers. (15 January, 2016). *Rachel Bloom Used to Be Seth's Intern - Late Night with Seth Meyers*. YouTube. https://www.youtube.com/watch?v=aQYgdDFNBr0.

Levin, Beth. (1993). *English verb classes and alternations*. Chicago, IL: Chicago University Press.

Levinson, Stephen & Asifa Majid. (2014). Differential ineffability and the senses. *Mind and Language* 29(4), 407-427.

Lynott, Dermot & Louise Connell. (2009). Modality exclusivity norms for 423 object properties. *Behavior Research Methods* 41, 558-564.

Lynott, Dermot & Louise Connell. (2013). Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior Research Methods* 45(2), 516-526.

Mandelkern, Matthew. (2019). What 'must' adds. *Linguistics and Philosophy* 42, 225-266.

Majid, Asifa & Niclas Burenhult. (2014). Odors are expressible in language, as long as you speak the right language. *Cognition* 130(2), 266-270.

Majid, Asifa, Seán G. Roberts, Ludy Cilissen, Karen Emmorey, Brenda Nicodemus, Lucinda O'Grady, Bencie Woll, Barbara LeLan, Hilário de Sousa, Brian L. Cansler, Shakila Shayan, Connie de Vos, Gunter Senft, N. J. Enfield, Rogayah A. Razak, Sebastian Fedden, Sylvia Tufvesson, Mark Dingemanse, Ozge Ozturk, Penelope Brown, Clair Hill, Olivier Le Geun, Vincent Hirtzel, Rik van Gijn, Mark A. Sicoli & Stephen C. Levinson. (2018). Differential coding of perception in the world's languages. *Proceedings of the National Academy of Sciences* 115(45), 11369-11376.

Matthewson, Lisa. (2012). Evidence about evidentials: Where fieldwork meets theory. In Britta Stolterfoht & Sam Featherston (eds.), *Empirical approaches to linguistic theory: Studies in meaning and structure* (Studies in Generative Grammar 111): 85-114. Berlin, DE & Boston, MA: De Gruyter Mouton.

McNally, Louise & Isidora Stojanovic. (2017). Aesthetic adjectives. In James O. Young (ed.), *Semantics of aesthetic judgments*: 17-37. Oxford, UK: Oxford University Press.

Muñoz, Patrick J. (2019). *On tongues: The grammar of experiential evaluation*. Chicago, IL: University of Chicago dissertation.

Murray, Sarah E. (2017). *The semantics of evidentials*. Oxford, UK: Oxford University Press.

Ortony, Andrew. (1979). Beyond literal similarity. *Psychological Review* 86(3), 161-180.

Osgood, Charles E., George J. Suci & Percy H. Tannenbaum. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.

Petersen, Wiebke, Jens Fleischhauer, Hakan Beseoglu & Peter Bücker. (2007). A frame-based analysis of synaesthetic metaphors. *Baltic International Yearbook of Cognition, Logic and Communication* 3(1), 1-22.

Petersen, Wiebke & Thomas Gamerschlag. (2014). Why chocolate eggs can taste old but not oval: A frame-theoretic analysis of inferential evidentials. In Thomas Gamerschlag, Doris Gerland, Rainer Osswald & Wiebke Petersen (eds.), *Frames and concept types: Applications in language and philosophy*: 199–220. Dordrecht, NL: Springer.

Pogue, Amanda & Michael K. Tanenhaus. (2018). Learning from uncertainty: Exploring and manipulating the role of uncertainty on expression production and interpretation. *Proceedings of the Annual Meeting of the Cognitive Science Society* 40, 2266–2271.

Poortvliet, Marjolein. (2018). *Perception and predication: A synchronic and diachronic analysis of Dutch descriptive perception verbs as evidential copular verbs*. Oxford, UK: University of Oxford dissertation.

Popova, Yanna. (2005). Image schemas and verbal synaesthesia. *From Perception to Meaning: Image Schemas in Cognitive Linguistics* 29, 395-419.

Porat, Roy & Yeshayahu Shen. (2015). Imposed metaphoricity. *Metaphor and Symbol* 30, 77-94.

Porat, Roy & Yeshayahu Shen. (2017). Metaphor: the journey from bidirectionality to unidirectionality. *Poetics Today* 38(1), 123–140.

Postal, Paul M. (1971). *Cross-over phenomena*. New York, NY: Holt, Rinehart and Winston.

Potsdam, Eric & Jeffery Runner. (2001). Richard returns: Copy-raising and its implications. In Mary Andronis, Christopher Ball, Heidi Elston & Sylvian Neuvel (eds.), *Papers from the 37th Regional Meeting of the Chicago Linguistic Society*: 453–468. Chicago, IL: Chicago Linguistic Society.

R Development Core Team. (2020). *R: a language and environment for statistical computing*. Vienna, AUST: R Foundation for Statistical Computing.

Roberts, Craige. (2004). Information structure in discourse. *Semantics and Pragmatics* 5, 1-69.

Rogers, Andrew D. Jr. (1974). *Physical perception verbs in English: A study in lexical relatedness*. Los Angeles, CA: University of California, Los Angeles dissertation.

Ronga, Irene, Carla Bazzanella, Ferdinando Rossi & Giandomenico Iannetti. (2012). Linguistic synaesthesia, perceptual synaesthesia, and the interaction between multiple sensory modalities. *Pragmatics and Cognition* 20(1), 135-167.

San Roque, Lila, Kobin H. Kendrick, Elisabeth Norcliffe, Penelope Brown, Rebecca Defina, Mark Dingemanse, Tyko Dirksmeyer, NJ Enfield, Simeon Floyd, Jeremy Hammond, Giovanni Rossi, Sylvia Tufvesson, Saskia van Putten & Asifa Majid. (2015). Vision verbs dominate in conversation across cultures, but the ranking of non-visual verbs varies. *Cognitive Linguistics* 26(1), 31-60.

Sassoon, Galit W. (2013). A typology of multidimensional adjectives. *Journal of Semantics* 30(3), 335-380.

Scontras, Gregory, Michael Henry Tessler & Michael Franke. (2018). *Probabilistic language understanding: An introduction to the Rational Speech Act framework*. https://www.problang.org.

Shen, Yeshayahu (1997). Cognitive constraints on poetic figures. *Cognitive Linguistics* 8(1), 33-71.

Shen, Yeshayahu & Michal Cohen. (1998). How come silence is sweet but sweetness is not silent: a cognitive account of directionality in poetic synaesthesia. *Language and Literature* 7(2), 123-140.

Shen, Yeshayahu & Ravid Eisenman. (2008). Heard melodies are sweet, but those unheard are sweeter: synaesthesia and cognition. *Language and Literature* 17(2), 101-121.

Shen, Yeshayahu. & David Gil. (2008). Sweet fragrances from Indonesia: a universal principle governing directionality in synaesthetic metaphors. In Willie van Peer, & Jan Auracher (eds.), *New beginnings in literary studies*: 49-71. Newcastle, UK: Cambridge Scholars Publishing.

Shen, Yeshayahu & Osnat Gadir. (2009). How to interpret the music of caressing: target and source assignment in synaesthetic genitive constructions. *Journal of Pragmatics* 41(2), 357-371.

Shibuya, Yoshikata, Hajime Nozawa & Toshiyuki Kanamaru. (2007). Understanding synesthetic expressions: vision and olfaction with the physiological = psychological model. In Martina Plümacher & Peter Holz (eds.), *Speaking of colors and odors*: 203–226. Amsterdam, NL: John Benjamins Publishing Company.

Shinohara, Kazuko & Akira Nakayama. (2011). Modalities and directions in synaesthetic metaphors in Japanese. *Cognitive Studies: Bulletin of the Japanese Cognitive Science Society* 18(3), 491-507.

Simons, Mandy. (2007). Observations on embedding verbs, evidentiality, and presupposition. *Lingua* 6, 1034-1056.

Solt, Stephanie. (2018). Multidimensionality, subjectivity and scales: Experimental evidence. In Elena Castroviejo, Louise McNally & Galit W. Sassoon (eds.), *The Semantics of gradability, vagueness, and scale structure: Experimental perspectives* (Language, Cognition, and Mind 4): 59-91. Cham, CH: Springer.

Staniewski, Przemysław & Adam Gołębiowski. (2021). To what extent can source-based olfactory verbs be classified as copulas? In Łukasz Jędrzejowski & Przemysław Staniewski (eds.), *The linguistics of olfaction*: 403-448. Amsterdam, NL: John Benjamins Publishing Company.

Steen, Gerard, Lettie Dorst, Berenike Herrmann, Anna Kaal, Tina Krennmayr & Trijntje Pasma. (2010). *A method for linguistic metaphor identification: From MIP to MIPVU*. Amsterdam, NL: John Benjamins Publishing Company.

Strik Lievers, Francesca. (2015). Synaesthesia: A corpus-based study of cross-modal directionality. *Functions of Language* 22(1), 69-95.

Strik Lievers, Francesca & Winter, Bodo. (2018). Sensory language across lexical categories. *Lingua* 204, 45-61.

Taniguchi, Kazumi. (1997). On the semantics and development of copulative perception verbs in English: A cognitive perspective. *English Linguistics* 14, 270-299.

Ullmann, Stephen. (1945). Romanticism and synaesthesia: a comparative study of sense transfer in Byron and Keats. *Publications of the Modern Language Association of America* 60, 811-827.

Ullmann, Stephen. (1957). *The principles of semantics*. Oxford, UK: Blackwell.

Ünal, Ercenur & Anna Papafragou. (2020). Relations between language and cognition: Evidentiality and sources of knowledge. *Topics in Cognitive Science* 12, 115-135.

Usoniene, Aurelia. (2000). On the modality of the English verbs of seeming. *Belgian Journal of Linguistics* 14(1), 185-205.

Verhagen, Justus V. & Lena Engelen. (2006). The neurocognitive bases of human multimodal food perception: sensory integration. *Neuroscience and Biobehavioral Reviews* 30, 613-650.

Viberg, Åke. (1983). The verbs of perception: A typological study. *Linguistics* 21(1), 123-162.

Viberg, Åke. (2019). Phenomenon-based perception verbs in Swedish from a typological and contrastive perspective. *Syntaxe et Sémantique* 20, 17-48.

Werning, Markus, Jens Fleischhauer & Hakan Beseoglu. (2006). The cognitive accessibility of synesthetic metaphors. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, 2365-2370.

Whitt, Richard J. (2009). Auditory evidentiality in English and German: The case of perception verbs. *Lingua* 119, 1083-1095.

Whitt, Richard J. (2010). *Evidentiality and perception verbs in English and German*. Oxford, UK: Peter Lang.

Whitt, Richard. J. (2011). (Inter)Subjectivity and evidential perception verbs in English and German. *Journal of Pragmatics* 43, 347-360.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Stephan Milton Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Paige Seide, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo & Hiroaki Yutaniet. (2019). Welcome to the Tidyverse. *Journal of Open Source Software* 4(43), 1686.

Williams, Joseph M. (1976). Synaesthetic adjectives: a possible law of semantic change. *Language* 52(2), 461-478.

Winkler, Alissa D., Lothar Spillman, John S. Werner & Michael A. Webster. (2015). Asymmetries in blue-yellow color perception and in the color of 'the dress'. *Current Biology* 25(13), R547-R548.

Winter, Bodo. (2016). *The sensory structure of the English lexicon*. Merced, CA: University of California, Merced dissertation.

Winter, Bodo. (2019a). *Sensory linguistics: language, perception, and metaphor*. Amsterdam, NL: John Benjamins Publishing Company.

Winter, Bodo. (2019b). Synaesthetic metaphors are neither synaesthetic nor metaphorical. In Laura J. Speed, Asifa Majid, Lila San Roque & Carolyn O'Meara (eds.), *Perception metaphors*: 105-126. Amsterdam, NL: John Benjamins Publishing Company.

Wise, Hilary. (1997). *The vocabulary of modern French: origins, structure and function*. London, UK: Routledge.

Wolff, Phillip & Dedre Gentner. (2011). Structure-mapping in metaphor comprehension. *Cognitive Science* 35(8), 1456-1488.

Yu, Ning. (2003). Synaesthetic metaphor: a cognitive perspective. *Journal of Literary Semantics* 32(1), 19-34.

תקציר

עבודת הדוקטורט עוסקת בשפה חושית, כלומר בביטויים לשוניים המשמשים דוברים לתקשורת אודות תפיסה ורשמים חושיים. העבודה מציבה את השפה החושית הן כמושא מחקר בפני עצמה, והן כאמצעי לחקירה של תופעות לשוניות החופפות את השפה החושית. העבודה מתמקדת בתת-קבוצה של ביטויים חושיים אותם אני מכנה פעלי חישה אוגדיים (copulative perception verbs), כגון *נראה* ו-*נשמע*.

פעלי חישה אוגדיים זכו לתשומת לב מועטה יחסית לביטויים חושיים אחרים, במיוחד בשפות שמקורן לא באירופה. על כן, המטרה הראשונה של עבודת הדוקטורט היא לקדם את הידע הבלשני אודות פעלים אלו, באמצעות חקירה אמפירית מעמיקה של השימוש בהם בשפה העברית. המטרה השנייה היא לגייס את המאפיינים הייחודיים של פעלים אלו על מנת לתת מענה לשתי שאלות תאורטיות פתוחות במחקר הבלשני: האחת בנושא אבידנציאליות (evidentiality), קרי דיווח על מקורות המידע העומדים לרשות הדובר, והשנייה בנושא מטפורות סינסתטיות (synesthetic metaphors), קרי ביטויים מטפוריים בהם חל מיפוי של תכונות מֵאופנוּת חושית (sensory modality) אחת לאופנוּת חושית אחרת.

כל אחד משלושת פרקי עבודת הדוקטורט מהווה פרויקט מחקר בלשני העומד בפני עצמו. בכל פרק נסקרת ספרות הנוגעת לתופעה לשונית, מזוהה שאלת מחקר, מתוארים הניבויים של תאוריות שונות, ונאספים נתונים אמפיריים במטרה להכריע אילו ניבויים אכן מתקיימים. לכל אורכה, עבודת הדוקטורט לא מוגבלת למסגרת תאורטית יחידה, אלא מתכתבת עם מגוון גישות, ובכללן סמנטיקה פורמאלית, בלשנות קוגניטיבית, בלשנות מבוססת-שימוש, ופסיכו-בלשנות. כמו כן נעשה שימוש במגוון כלים מתודולוגיים, ובכללם ניתוחי קורפוס, ניסויים פסיכו-בלשניים, ומודלים חישוביים.

הפרק הראשון חוקר את התבניות השונות שבהן מופיעים פעלי חישה אוגדיים בעברית. בספרות על פעלי חישה, האבחנה הבסיסית היא בין פעלים המורים על חוויית תפיסה חושית, שהנושא הדקדוקי שלהם הוא הסובייקט התופס דבר מה בחושיו (כגון *ראה* ו-*שמע*), לבין פעלים המורים על רשמים חושיים, שהנושא הדקדוקי שלהם הוא האובייקט הנתפס (Viberg 1983). את האחרונים ניתן לסווג כפעלים שמהווים פרדיקָטים בפני עצמם (כגון *מבהיק* ו-*מרשרש*), או כפעלים אוגדיים, אשר דורשים פרדיקט או פסוקית בתור משלים על מנת להתפרש בתור פרדיקטים (כגון *נראה* ו-*נשמע*).

מבחינה סמנטית, הספרות מבחינה בין שני שימושים של פעלי חישה אוגדיים. בשימושים ייחוסיים (attributary), הפעל מייחס את המשלים לרושם חושי מסוים של הנושא הדקדוקי, ואילו בשימושים ממסגרים (parenthetical), הפעל מאייך טענה המורכבת מהנושא הדקדוקי והמשלים (Gisborne 2010). לדוגמא, המבע *Rachel sounds weird* יכול להתפרש כתיאור הקול (הרושם השמיעתי) של רייצ'ל, או כמאייך את הטענה 'רייצ'ל מוזרה'. אין בספרות הסכמה רחבה בשאלה האם השימושים הממסגרים הינם

ביסודם אפיסטמיים, קרי מדווחים על מידת האמונה של הדובר בטענה, או אבידנציאליים, קרי מדווחים את מקור המידע שמוביל את הדובר לטענה, או שמא הם אפיסטמיים ואבידנציאליים גם יחד. יתרה מכך, בשפות שבהן פעלי חישה אוגדיים נחקרו בעבר, לא נמצאה שיטה מהימנה לזיהוי כל אחד מהשימושים.

בעברית, בניגוד לשפות שבהן פעלי חישה אוגדיים נחקרו עד כה, פעלים אלו יכולים לקחת בתור משלים הן תואר פועל (כך ב-*נשמעת מוזר* ו-*נראים נחמד*) והן שם תואר (כך ב-*נשמעת מוזרה* ו-*נראים נחמדים*). לצד שתי התבניות האלו, עברית מאפשרת גם תבנית סתמית ללא נושא דקדוקי (כגון *נראה שירד גשם*). נשאלת השאלה, האם שלוש התבניות נבדלות זו מזו מבחינה סמנטית? במענה לשאלה זו, ערכתי מחקר קורפוס מקיף על המופעים של שלוש התבניות באתרי אינטרנט (כ-100,000 מופעים בסך הכל), והשוויתי ביניהן בשיטת ניתוח סטטיסטי ייעודית לזיהוי הבחנות בין תבניות לשוניות דומות (Gries & Stefanowitsch 2010).

הממצאים המרכזיים העולים מניתוח הקורפוס נוגעים להשוואה בין התבנית שבה תואר פועל נלקח כמשלים, לבין התבנית בה שם תואר נלקח כמשלים. במופעים של התבנית הראשונה, המשלים נטה להיות אחד מקבוצה מצומצמת של פרדיקטים רב-מימדיים וסובייקטיביים, ללא תלות בחוש (כגון *נהדר*, *טוב*, *אחרת* ו-*מוזר*). הנושא הדקדוקי נטה להיות אובייקט הניתן לתפיסה בחוש הרלוונטי (כגון *אנשים נראים*, *מוזיקה נשמעת*, ו-*חדר מריח*). לעומת זאת, במופעים של התבנית השנייה, בה שם תואר נלקח כמשלים, המשלים נטה להיות פרדיקט מופשט, שלא בהכרח ניתן לייחוס לרושם החושי הרלוונטי (כגון *נראה הגיוני*, *נשמע מופרך*, ו-*מריח מפוקפק*). יתר על כן, הנושא הדקדוקי נטה להיות אובייקט מופשט, שלא בהכרח ניתן לתפיסה בחוש הרלוונטי (כגון *שאלה נראית*, *צורה נשמעת*, ו-*סיפור מריח*).

מחקר הקורפוס, אם כך, מצביע על הבדל סמנטי מובהק בין שתי התבניות. התבנית בה תואר פועל נלקח כמשלים מקיימת את המגבלות הצפויות לשימוש הייחוסי: הנושא הדקדוקי ניתן לתפיסה בחוש הרלוונטי, והמשלים ניתן לייחוס לרושם החושי הרלוונטי. למעשה, המשלימים של התבנית בעברית מוגבלים אף יותר מהצפוי, שכן אין בנמצא מופעים עם פרדיקטים חד-מימדיים (כגון *נראית חזק*, *נשמעים גבוה*, או *מרגישים חלק*). אי לכך, אני מציע ניתוח חדש לשימושים ייחוסיים בעברית, לפיו הפעל מגביל את הפירוש של המשלים לתת-קבוצה של מימדים הנתפסים בחוש מסוים, בדומה לניתוח שהוצע לביטויים כגון *מבחינת מראה* (Sassoon 2013).

התבנית בה שם תואר נלקח כמשלים מפרה את כל המגבלות הצפויות לשימוש הייחוסי, ומכאן שהיא דומה יותר לשימושים הממסגרים. אולם ממצאי הקורפוס לא מאפשרים לקבוע האם המשמעות הבסיסית של התבנית היא אפיסטמית או אבידנציאלית. כך גם לגבי התבנית הסתמית. על מנת לזהות האם מתקיימת הבחנה סמנטית בין שתי התבניות הנותרות, יצרתי וערכתי ניסוי העדפת משפטים. בניסוי, לנבדקים מוצגים

הקשרים שיחיים שונים, והם נדרשים לבחור איזו מבין שתי התבניות עדיפה בכל הקשר. כל אחד מההקשרים סותר את אחת המשמעויות שהוצעו בספרות. בהקשר שסותר את המשמעות האפיסטמית, הדובר מאמין כי הטענה שאותה מאייך הפעל היא שקרית. בהקשר שסותר את המשמעות האבידנציאלית, מקור המידע של הדובר לטענה אינו תפיסה ישירה של הנושא הדקדוקי. להלן דוגמאות לבחירות שניתנו לנבדקים בניסוי:

1. אני יודע שהמלצרית צעירה, אבל...
   א) היא נראית לי מבוגרת.
   ב) נראה לי שהיא מבוגרת.
2. מעולם לא ראיתי את המלצרית, אבל...
   א) היא נראית לי מבוגרת.
   ב) נראה לי שהיא מבוגרת.

תוצאות הניסוי מצביעות בבירור על הבדל בין שתי התבניות הנחקרות. התבנית בה נלקח שם תואר כמשלים עדיפה בהקשרים שסותרים את המשמעות האפיסטמית (1א בדוגמא לעיל), בעוד שהתבנית הסתמית עדיפה בהקשרים שסותרים את המשמעות האבידנציאלית (2ב בדוגמא לעיל). המסקנה המתבקשת מכך היא שהתבנית הסתמית מקודדת משמעות אפיסטמית, ואילו התבנית בה נלקח שם תואר כמשלים מקודדת משמעות אבידנציאלית. לאור זאת, המחקר מעלה שלכל אחת משלוש התבניות בהן מופיעים פעלי חישה אוגדיים בעברית ישנה פונקציה סמנטית ייחודית לה.

כזכור, כל אחד מפרקי עבודת הדוקטורט מהווה פרויקט מחקר העומד בפני עצמו, הנוגע לפעלי חישה אוגדיים. הפרק השני חוקר סוגיה פרגמטית הנוגעת לאבידנציאליות. ביטויים אבידנציאליים הם כאלה שבאמצעותם דובר מדווח על מקור המידע שמעיד על טענה מסוימת. כפי שנטען בפרק הראשון של הדוקטורט, לפעלי חישה אוגדיים כמו *נראה* ו-*נשמע* יש גם שימוש אבידנציאלי. ידוע כי השימוש בביטויים אבידנציאליים נוטה לייצר היסק פרגמטי, לפיו הדובר אינו לחלוטין משוכנע באמיתות הטענה שעליה המידע מעיד. כך, דובר המשתמש במבע *השמלה נראית חדשה* נשפט על ידי מאזינים כמהוסס לגבי אמיתות הטענה 'השמלה חדשה', במיוחד בהשוואה לדובר המשתמש במבע העירום (bare) *השמלה חדשה* (Pogue & Tanenhaus 2018). מדובר בהיסק פרגמטי, שכן הוא בר-ביטול ולא עולה בהכרח בכל הקשר שיחי.

בספרות הועלו שני הסברים נפרדים להיסק האי-ודאות שנקשר בביטויים אבידנציאליים. האחד נשען על הקשה חוץ-לשונית אודות סוג המידע עליו הביטוי האבידנציאלי מדווח, ובעיקר אודות מידת הישירות של העדות (von Fintel & Gillies 2010). למשל, במבע *השמלה נראית חדשה* הדובר מדווח שהמידע העומד לרשותו הוא מראה השמלה. היות ומראה משמש עדות בלתי ישירה לטענה 'השמלה חדשה', מתקבל על

הדעת שהדובר אינו משוכנע באמיתות הטענה. על פי ההסבר החוץ-לשוני, מאזינים מייחסים לדובר מידת ודאות פרופורציונית למידת הישירות של המידע העומד לרשותו.

ההסבר השני נשען על הקשה גרייסיאנית, קרי הקשה פרגמטית המבוססת על עקרון שיתוף הפעולה של גרייס (Grice 1989). בהקשה גרייסיאנית, המאזינים שוקלים את מבע הדובר כנגד מבעים אלטרנטיביים שבהם הדובר יכול היה להשתמש, אך נמנע מלעשות כן. למשל, דובר שהשתמש במבע *השמלה נראית חדשה* יכול היה להשתמש תחת זאת במבע העירום, הקצר יותר, *השמלה חדשה*. על בסיס עקרון שיתוף הפעולה, המאזינים מסיקים שהייתה לדובר סיבה תקשורתית לבחור במבע הארוך יותר. סיבה אפשרית היא לאותת למאזינים שהדובר לא משוכנע באמיתות הטענה 'השמלה חדשה' (Degen et al. 2015).

שני ההסברים לעיל, החוץ-לשוני והגרייסיאני, מייצרים ניבויים שונים אודות קבוצת מבעים שלא נחקרה עד כה, אשר מכילים ביטוי אבידנציאלי המדווח על מידע ישיר במידה המירבית. למשל, במבע *השמלה נראית כחולה* הדובר מדווח שהמידע העומד לרשותו הוא מראה השמלה, והרי זוהי העדות הישירה ביותר שתיתכן לטענה אודות צבע השמלה. מכאן שעל פי ההסבר החוץ-לשוני, המאזינים צפויים לייחס לדובר מידת ודאות גבוהה. מנגד, על פי ההסבר הגרייסיאני, המאזינים צפויים לשקול את המבע כנגד המבע העירום, ולפיכך עשויים למרות הכל להקיש שהדובר לא משוכנע באמיתות הטענה.

אני טוען כי שני ההסברים אינם סותרים זה את זה, ומציע לשלבם לכדי הסבר "היברידי" אחד שמערב הן הקשה חוץ-לשונית והן הקשה גרייסיאנית. על פי הסבר זה, המאזינים שוקלים כל מבע עם ביטוי אבידנציאלי כנגד המבע העירום, ולכן מסיקים אי-ודאות מצד הדובר כברירת מחדל. אם הביטוי האבידנציאלי מדווח על מידע בלתי ישיר, היסק האי-ודאות מתיישב עם ההקשה החוץ-לשונית לגבי הקשר בין ישירות המידע לודאות הדובר. אולם אם הביטוי האבידנציאלי מדווח על מידע ישיר, כמו במבע *השמלה נראית כחולה*, היסק האי-ודאות סותר את ההקשה החוץ-לשונית. המאזינים נדרשים למהלך היקשי נוסף על מנת ליישב את הסתירה. אפשרות אחת היא להסיק שהמידע שעומד לרשות הדובר, על אף היותו ישיר, אינו מהימן, למשל אם השמלה נחזתה מבעד לזכוכית צבועה.

במטרה לבחון את ניבויי שלושת ההסברים, יצרתי וערכתי שני ניסויים. בראשון, הנבדקים משמשים בתפקיד מאזינים. מוצגים להם מבעים שונים, חלקם עם ביטוי אבידנציאלי וחלקם עירומים, והם נדרשים לדרג את מידת הודאות של הדובר לגבי הטענה. בשונה מניסויים דומים שנערכו בעבר, שבהם נחקרו רק ביטויים אבידנציאליים שמדווחים על מידע בלתי ישיר (Degen et al. 2019), בניסוי זה נכללו ביטויים אבידנציאלים שמדווחים על מידע במנעד של ישירות, כולל גם מידע ישיר במידה המירבית. תוצאות הניסוי מראות שמאזינים מסיקים אי-ודאות מכל שימוש בביטוי אבידנציאלי, בין אם הוא מדווח על מידע ישיר או על מידע בלתי ישיר. בנוסף, מידת האי-ודאות תלויה במידת הישירות של המידע, אך התלות אינה

ליניארית. הן מידע בלתי ישיר (כגון *השמלה נראית שמנונית*) והן מידע ישיר במידה המירבית (כגון *השמלה נראית כחולה*) מובילים מאזינים להסיק מידה גבוהה יותר של אי-ודאות, מאשר מידע ישיר במידה בינונית (כגון *השמלה נראית נקייה*).

בניסוי השני, הנבדקים משמשים בתפקיד דוברים. מוצגים להם הקשרים שיחיים שונים, והם נדרשים לבחור האם מבע עירום או מבע עם ביטוי אבידנציאלי עדיף בכל הקשר. כל הקשר מתאר נסיבות שבהן עומד לרשות הדובר מידע ישיר או בלתי ישיר (לדוגמא, חלון שניתן להשקיף דרכו או דלת שניתן לצותת דרכה), וכן מהימן או בלתי מהימן (לדוגמא, חלון שקוף או חלון צבוע). תוצאות הניסוי מראות שבנסיבות בהן המידע בלתי מהימן, נכונות הדוברים להשתמש בביטוי אבידנציאלי עולה עם ישירות המידע.

כדי להמחיש שתוצאות שני הניסויים תומכות בהסבר ההיברידי, פיתחתי מודל חישובי של ההסבר במסגרת הידועה בשם "פעולת דיבור רציונלית" (Frank & Goodman 2012). במודל מסוג זה, דוברים ומאזינים מיוצגים בתור פונקציות מתמטיות, שמשתמשות בכלים מתחום הסטטיסטיקה הבייזיאנית על מנת לייצר ולפרש מבעים. החידוש במודל שלי, לעומת מודלים קיימים, הוא בייצוג של ישירות ומהימנות המידע העומד לרשות הדובר. ישירות ומהימנות המידע מוזנים כקלט לפונקציה שמייצגת את הדובר ומשפיעים על המבע שהוא מייצר כפלט. אותו מבע מוזן כקלט לפונקציה שמייצגת את המאזין, אשר מוציא כפלט היסקים אודות מידת הודאות של הדובר, כמו גם אודות ישירות ומהימנות המידע שעומד לרשותו.

לבסוף, הפרק השלישי של הדוקטורט מגייס תבנית מיוחדת של פעלי חישה אוגדיים על מנת לחקור את התופעה של העדפות כיווניות במטפורות סינסתטיות. מטפורה סינסתטית היא תיאור של גירוי חושי במונחים של אופנות חושית אחרת. למשל, *צרימה דוקרת* היא תיאור של גירוי שמיעתי באמצעות מונח מישושי (tactile). מדובר במטפורה כיוון שתיאור כזה מהווה מיפוי של תכונות מתחום קוגניטיבי אחד לאחר. הספרות מראה שבמטפורות סינסתטיות ישנן העדפות כיווניות, כלומר מיפוי מחוש א' לחוש ב' עדיף מכל מיני בחינות על פני המיפוי ההפוך, מחוש ב' לחוש א'. לדוגמא, מיפויים מתחום המישוש לתחום השמיעה, כגון *צרימה דוקרת*, נפוצים יותר בדיבור ובשירה מאשר מיפויים הפוכים, משמיעה למישוש, כגון *דקירה צורמת* (Ullmann 1957, Shen 1997). העמדה הרווחת בספרות היא כי מכלול ההעדפות הכיווניות משקף היררכיה של החושים, שבראשה ראייה ושמיעה ובתחתיתה מישוש וטעם, כך שמיפויים במעלה ההיררכיה עדיפים על פני מיפויים במורדה.

מגוון רחב של גורמים הוצעו בספרות כתורמים להעדפות הכיווניות במטפורות סינסתטיות. אני מסווג אותם כגורמים לקסיקליים או גורמים חושיים. גורמים לקסיקליים קשורים בביטויים הלשוניים הספציפיים שמופיעים במטפורה, וכוללים בין היתר שכיחות, ריגושיות (affectivity), ודירוגיות (gradability). ניסויים

וניתוחי קורפוס קודמים מצאו שלגורמים אלו ישנה השפעה מובהקת, חוצת אופנויות חושיות, על העדפות כיווניות (Petersen et al. 2007, Winter 2016).

מנגד, גורמים חושיים אינם תלויים בביטויים הלשוניים בהם נעשה שימוש, אלא נובעים ישירות מהבדלים בין האופנויות החושיות עצמן. לדוגמא, הוצע שמגע ישיר בין איבר החישה לגירוי החושי, או היעדרו, הוא אחד הגורמים להעדפות כיווניות. אולם ההשפעה של גורמים חושיים מעולם לא הוכחה אמפירית, כיוון שבכל מחקר שבחן אותה הייתה השפעה מתערבת של גורמים לקסיקליים (Winter 2019a).

הסיבה לכך, לטענתי, היא שמחקרים קודמים התמקדו במטפורות סינסתטיות שבהן האופנויות החושיות נתרמות על ידי שמות עצם ושמות תואר, כגון *צרימה דוקרת*. במטפורות מעין אלה בלתי אפשרי לשלוט בכל הגורמים הלקסיקליים, שכן שמות עצם ושמות תואר הקשורים בחושים שונים אינם ברי השוואה. לעומת זאת, פעלי חישה אוגדיים מאפשרים שליטה כזאת, היות ומדובר בקבוצה סגורה של ביטויים עם משמעויות מקבילות. השימוש הייחוסי של הפעלים מאפשר יצירת מטפורות סינסתטיות בתבנית אנלוגיה, שבה גורמים כגון ריגושיות ודירוגיות מיטשטשים. לדוגמא, באנלוגיה *התמונה נראית כמו שהמוזיקה נשמעת*, חל מיפוי מתחום השמיעה לתחום הראייה, אך התכונות הממופות נותרות מרומזות.

במטרה לבחון האם העדפות כיווניות מתגלות גם בהיעדר גורמים לקסיקליים, כלומר אך ורק בהשפעת גורמים חושיים, יצרתי וערכתי ניסוי דירוג טבעיות. בניסוי, לנבדקים מוצגים מטפורות סינסתטיות שונות בתבנית האנלוגיה לעיל, והם נדרשים לדרג עד כמה כל מטפורה טבעית, קבילה וניתנת לפירוש. המטפורות שהוצגו מכסות את כל המיפויים האפשריים בין חמשת החושים, בשני הכיוונים האפשריים.

תוצאות הניסוי תומכות בטענה שישנם גורמים חושיים המשפיעים באופן ישיר על העדפות כיווניות במטפורות סינסתטיות. עם זאת, העדפות כיווניות התגלו רק בצמדי חושים מסוימים. בפרט, מיפויים ממישוש לשמיעה עדיפים על המיפויים ההפוכים משמיעה למישוש, ומיפויים מראייה לשמיעה עדיפים על המיפויים ההפוכים משמיעה לראייה. בניגוד לסברה הרווחת, לא נמצאה עדות לטענה שמכלול ההעדפות הכיווניות משקף היררכיה של החושים. איששתי את ממצאי הניסוי גם במחקר קורפוס בהיקף מצומצם (10,000 מבני אנלוגיה). לאור ממצאים אלו, אני מעלה את הסברה שהיררכיית החושים המקובלת על מרבית החוקרים בתחום היא ממצא שווא. ייתכן כי מכלול ההעדפות הכיווניות שהתגלה בעבר נובע מבליל של גורמים חושיים ולקסיקליים בלתי תלויים, שרק במקרה מצטברים לכדי מה שנראה כמו היררכיה אוניברסלית יחידה.

אוניברסיטת תל-אביב

הפקולטה למדעי הרוח ע"ש לסטר וסאלי אנטין

ביה"ס לפילוסופיה, בלשנות ולימודי מדע

פעלי חישה אוגדיים: אבידנציאליות וסינסתזיה

חיבור זה הוגש לשם קבלת התואר "דוקטור לפילוסופיה"

באוניברסיטת תל אביב

על ידי

אלון פישמן

העבודה הוכנה בהדרכת:

מירה אריאל

ישעיהו שן

תאריך: יוני 2023