Projection variability in Paraguayan Guaraní

September 12, 2019

Abstract

Projective content is heterogeneous, with classes of projective content differing in several properties (e.g., Potts 2005; Tonhauser et al. 2013). Recently, Tonhauser et al. (2018) found that projective content in English varies in its projectivity both between and within classes, and also that there is by-participant and by-lexical content projection variability. This paper shows that projection variability is not unique to English but also attested in Paraguayan Guaraní, a Tupí-Guaraní language that is genetically unrelated to and typologically different from English. This finding suggests that projection variability may be a cross-linguistically universal property of projective content. The comparison of English and Paraguayan Guaraní also reveals parallels in how projective the content associated with a translation pair is. This finding strengthens the empirical support for the position that some projective content is nondetachable (e.g.,Levinson and Annamalai 1992; Simons 2001; Abrusán 2011, 2016; Tonhauser et al. 2013). The paper discusses implications for analyses of projective content, which differ in whether they lead us to expect projection variability and cross-linguistic similarities in projection variability. The paper also addresses methodological considerations in exploring projection variability in fieldwork-based research.

1 Introduction

Projective content is utterance content to whose truth the speaker may be taken to be committed even when the expression associated with the content is embedded under an entailment-canceling operator, such as negation or the polar question operator (Karttunen 1971; Chierchia and McConnell-Ginet 1990; Potts 2005, i.a.). Two well-known projective contents are the content of the complement of *discover* in (1) and the content of the non-restrictive relative clause (NRRC) in (2). Both contents are considered projective because speakers may be taken to be committed to their truth regardless of whether speakers utter the affirmative sentences in (1a) and (2a) or the polar question variants in (1b) and (2b).

(1) a. Juana discovered that Carla was dancing.
b. Did Juana discover that Carla was dancing?

(2) a. Juana, who is an avid bicyclist, is fixing her bicycle.
b. Is Juana, who is an avid bicyclist, fixing her bicycle?

While projective content shares the property of being projective, it is otherwise heterogeneous. For instance, conventional implicatures, like the content of the NRRC in (2), differ from presuppositions, like the content of the complement of *discover* in (1), in being antibackgrounded and not plugged by verbs of saying (Potts 2005:§2.4.3). One goal of research on projective content is to determine which classes of projective content there are, i.e., which projective contents differ from one another in which empirical properties. Understanding empirical properties of projective content is essential to a second goal of research on projective content, which is to develop empirically adequate analyses of the different classes of projective content. For example, among conventional implicatures, distinct analyses have been given to supplements

and expressives because expressives, unlike supplements, code properties of speakers' emotional states (e.g., Potts 2005, 2007). And whereas the analyses in Heim 1983 and van der Sandt 1992 were intended to apply to all presuppositions, more recent analyses targeted subsets of presuppositions that were taken to be less projective than others (e.g., Abusch 2002, 2010; Abrusán 2011, 2016; Romoli 2015; Simons et al. 2017), where one content being less projective than another means that the speaker is less likely to be taken to be committed to the truth of the first than the second.

The assumption that projective content varies in its projectivity has been around for a long time (e.g., Karttunen 1971; Kadmon 2001; Potts 2005) but it has only recently been investigated systematically. The most comprehensive investigation to date is Tonhauser et al.'s 2018 experimental investigations of 19 projective contents of English: this work observed expected projection variability between classes of projective content but also unexpected projection variability within classes. For instance, their Experiment 1a found that presuppositions were overall less projective than conventional implicatures. Regarding between-class variability, their Experiment 1b found that the contents of the complements (CCs) of *be annoyed, discover* and *reveal*, which are all typically considered presuppositions, differed in projectivity such that the CC of *be annoyed* was more projective than that of *discover*, which in turn was more projective than that of *reveal*. As discussed in Tonhauser et al. 2018, the observed projection variability aligns neither with standard assumptions about relative projectivity nor with standardly-assumed distinctions among presuppositions, such as the distinction between 'hard' and 'soft' presupposition triggers or between 'factive' and 'semi-factive' predicates (e.g., Karttunen 1971; Kadmon 2001; Abusch 2010; Abrusán 2011, 2016). Further, the projection variability observed between presuppositions in English is not predicted by analyses currently on the market (Abusch 2002, 2010; Abrusán 2011, 2016; Romoli 2015; Simons et al. 2017). In sum, Tonhauser et al. (2018) showed that investigating projection variability is critical to developing empirically adequate analyses of projective content.

At this point, we have some insight into projection variability in English: in addition to Tonhauser et al. 2018, who investigated 19 contents, the projectivity of projective content was also investigated in Smith and Hall 2011 and Tonhauser et al. 2019. Smith and Hall 2011 investigated six contents, namely the existence content of cleft sentences, the content of the complement of *know*, the participation-content associated with *win*, the noun content of definite noun phrases and the contents of NRRCs and an expressive. They found that the projective contents associated with *win* and *know* are less projective than the noun content of definite noun phrases but, contrary to expectation, not less projective than the existential content of cleft sentences. Contrary to Tonhauser et al. 2018, they did not observe that the conventional implicatures are more projective than the presuppositions. It is an open question what this difference is due to because the two sets of experiments differed on a number of dimensions, including the entailment-canceling operators that the relevant expressions were embedded under, the response tasks, the expressions explored and the number of items per expression. Tonhauser et al. (2019) investigated the projectivity of three contents, namely the prejacent of repetitive *again* (that the relevant event has happened before), the pre-state content of *stop* and the prejacent of manner adverb sentences (as well as the influence of prosody on projection). They found that the prejacent of manner adverbs is less projective than the pre-state content of *stop* but also, contrary to expectation given that *again* is considered a 'hard' trigger, that the prejacent of *again* is less projective than the pre-state content of *stop*. In sum, we know that projective content in English varies in its projectivity, but the relative projectivity of projective content has not yet been fully investigated.

Empirically adequate analyses must account for projective content and projection variability in English and other languages. Two core questions about projection variability across languages are the following:

(3) Questions about projection variability across languages
   a. What is the typological profile of projection variability? Is it unique to English, a cross-linguistically universal property of projective content, or observed in some but not all languages?
   b. If projection variability is cross-linguistically attested, is projective content associated with a

translation pair similarly projective or are there systematic cross-linguistic differences?

At this point, these two questions are wide open because there are almost no experimental investigations of projection variability in languages other than English. The only exception is Xue and Onea's 2011 investigation of 4 contents in German, which found that the content of the complement of German *wissen* 'know' is less projective than the content of the complement of *erfahren* 'discover/find out', both of which are less projective than the relevant projective contents of sentences with *auch* 'too' (that a parallel event is contextually salient) and repetitive *wieder* 'again'. Although Xue and Onea's investigation had limited empirical scope, it already suggests, in response to the question in (3a), that there is projection variability not just in English but also in German. Interestingly, contrary to what Xue and Onea (2011) observed for German *wissen* 'know' and *erfahren* 'discover/find out', Tonhauser et al. 2018 observed that the content of the complement of English *know* is more projective than that of *discover*. It would be premature, however, to conclude that the two languages exhibit systematic differences because Xue and Onea's 2011 and Tonhauser et al.'s 2018 experiments differed on a number of dimensions, including the entailment-canceling operators that the relevant expressions were embedded under, the response tasks, the number of expressions explored and the number of items per expression. In short, the question in (3b) has not yet been addressed.

In this paper, both questions in (3) are addressed through investigations of projection variability in Paraguayan Guaraní, a Tupí-Guaraní language spoken in Paraguay. Section 2 presents the findings of three investigations of the projectivity of a total of 21 contents which show, in response to question (3a), that projection variability is not unique to the Germanic languages English and German, but is also found in Paraguayan Guaraní. The observation that there is projection variability in English and Paraguayan Guaraní, two genetically unrelated and typologically different languages,[1] suggests that projection variability may be a cross-linguistically universal property of projective content. Section 2 ends with a discussion of the implications of these findings for analyses of projective content currently on the market as well as methodological considerations in exploring projection variability with theoretically untrained native speakers.

Question (3b) is addressed in section 3: a comparison of projection variability in English and Paraguayan Guaraní is made possible by the fact that projection variability in Paraguayan Guaraní was investigated using similar methods as and on the basis of translations of English expressions examined in Tonhauser et al. 2018. This comparison reveals similarities in how projective the contents associated with translations pairs in the two languages are, thereby expanding on Tonhauser et al.'s 2013 finding that projective content associated with translation pairs in the two languages are strikingly similar in the properties exhibited. As discussed in section 3, these cross-linguistic parallels strengthen the empirical support for the position that at least some projective content is nondetachable, i.e., not lexically specified but derived from truth-conditional meaning (e.g.,Levinson and Annamalai 1992; Simons 2001; Abrusán 2011, 2016; Tonhauser et al. 2013). The paper concludes in section 4.

2 Utterance content in Paraguayan Guaraní varies in projectivity

This section addresses question (3a) on the basis of the findings of three investigations of projection variability in Paraguayan Guaraní: an investigation based on one-on-one elicitation (section 2.1), a pilot study (section 2.2) and an experiment (section 2.3). The 21 contents whose projectivity was investigated across the three investigations are given in Table 1 together with the expression that the content is associated with. The color coding, which is used throughout the paper,[2] identifies the standard classification of the projective content associated with the English translations, as elaborated in the following. The six expression/content

[1] Paraguayan Guaraní, which is spoken by about five million people in Paraguay, is a mildly polysynthetic, agglutinative and head-marking language with a split-S argument marking system. Word order is relatively free and influenced by information structure. For overviews of the grammar see, e.g., Gregores and Suárez 1967; Velázquez-Castillo 2004a; Estigarribia 2017.

[2] The colors were chosen to maximize accessibility for readers with color vision deficiencies and to allow for grayscale printing.

pairs in 1.-6. were also investigated in Tonhauser et al. 2013: as mentioned above, the content of NRRCs (1.) is standardly taken to be a conventional implicature (given in orange) whereas the other 5 contents (2.-6.) are standardly taken to be presuppositions (given in magenta), though 3. and 4. have also received non-presuppositional analyses (for discussion see, e.g., Horn 1996, 2002; Beaver and Clark 2008; Roberts 2011). The 14 expression/content pairs in 7.-20. have not previously been investigated. The English translations of the verbs in 7.-14. are typically considered factive verbs and the contents of their complements (CCs) are taken to be presuppositions. By contrast, the English translations of the verbs in 15.-20. are typically considered non-factive verbs and their CCs are not analyzed as presuppositions. The CCs of *–hechauka* 'show, reveal' (15.) *–'e* 'say' (16.), *–mombe'u* 'confess, tell' (17.) and *–hendu* 'hear' (18.), given in purple, were included because the CCs of their English translations have been argued to be projective despite being non-factive (see, e.g., Schlenker 2010; Anand and Hacquard 2014; Spector and Egré 2015). The CCs of *–mo'ã* 'think' (19.) and *–ha'arõ* 'hope' (20.) given in black, were hypothesized to not project, like their English counterparts and main clause content (21.).[3] Thus, the 21 expression/content pairs allow for an investigation of projection variability both within and between classes of content.

The findings about projection variability presented in this section are based on three investigations with theoretically untrained speakers. The findings of the initial investigation, presented in section 2.1, was based on the elicitation of responses from native speakers in one-on-one sessions (see, e.g., Matthewson 2004 and the papers in Bochnak and Matthewson 2015); this is what is typically meant with 'fieldwork'. Sections 2.2 and 2.3 present the findings of a pilot study and an experiment; both build on the experimental design developed for English in Tonhauser et al. 2018. As discussed in section 2.4, the three investigations differed on a number of dimensions: in, for instance, the number of items per expression/content pair, the number of participants from whom responses were elicited, whether participants were asked to give categorical responses or responses on a Likert scale, how much data was collected from each participant, whether the data can be (and were) subjected to statistical analysis and how tightly various factors were controlled for. Section 2.4 concludes with a reflection on methods for investigating projection variability in the field.

To empirically investigate projection variability, we have to be clear about what it means for a content, like the CC of *discover* in (1b) to be less projective than another, like the content of the NRRC in (2b).

(1b) Did Juana discover that Carla was dancing?

(2b) Is Juana, who is an avid bicyclist, fixing her bicycle?

According to Tonhauser et al. 2018:498f., there are at least two interpretations of what it means for projectivity to be a gradient property rather than a binary, categorical one. Both interpretations are compatible with the standard definition of projective content as utterance content to whose truth speakers may be taken to be even when the expression associated with the content is embedded under an entailment-canceling operator (see references in section 1). On a first interpretation, one content being less projective than another means that listeners (or readers) take speakers (or writers) to be less committed to the truth of the first content than to the truth of the second. On this interpretation, projectivity being a gradient property is a consequence of speaker commitment to the truth of a content being a gradient property: to say that the content of the NRRC is more projective than the CC of *discover* means that, overall, speakers are more committed to the truth of the content of the NRRC than to the truth of the CC of *discover*. On a second interpretation, one content being less projective than another reflects the probability with which listeners take the speaker to be committed to the truth of the content. On this interpretation, speaker commitment may be a binary, categorical property and projection variability arises from the listener's uncertainty about whether the speaker is committed: to

[3]English *see* and *hear* are taken to be factive when the complement describes an event that can be seen or heard, respectively, and non-factive/evidential otherwise. The CCs of *–hecha* 'see' and *–hendu* 'hear' are assigned to different classes here because *–hecha* 'see' was combined with complements that described an event that can be seen whereas *–hendu* 'hear' was combined with complements that described an event that cannot be heard.

	Expression/content pair	One-on-one elicitation (section 2.1)	Pilot study (section 2.2)	Experiment (section 2.3)
1.	Content of NRRCs	(✓)		✓
2.	Possession content of possessive noun phrases	(✓)		✓
3.	Prejacent of exclusive clitic *=nte* 'only'	(✓)		✓
4.	Polar content of proximal adverb *haimete* 'almost'	(✓)		✓
5.	Pre-state content of change-of-state expression	(✓)		✓
6.	CC of *–kuaa* 'know'	✓	✓	✓
7.	CC of *–kuaa-uka* (know-CAUS) 'point out'	✓		✓
8.	CC of *–juhu* 'discover'	✓	✓	✓
9.	CC of *–topa* 'discover'	✓		✓
10.	CC of *–hecha* 'see'	✓		✓
11.	CC of *–hecha-kuaa* (see-know) 'understand, realize'	✓	✓	✓
12.	CC of *–vy'a* 'be happy'	✓	✓	✓
13.	CC of *–guerotĩ* 'be ashamed'	✓	✓	
14.	CC of *–ñembyasy* 'regret'	✓	✓	
15.	CC of *–hecha-uka* (see-CAUS) 'show, reveal'	✓		✓
16.	CC of *–'e* 'say'	✓		✓
17.	CC of *–mombe'u* 'confess, tell'	✓	✓	✓
18.	CC of *–hendu* 'hear'	✓		✓
19.	CC of *–mo'ã* 'think'	✓	✓	✓
20.	CC of *–ha'arõ* 'hope'	✓	✓	
21.	Main clause content	(✓)		✓

Table 1: Contents whose projectivity was investigated, with the associated expression. '✓' indicates that the expression/content pair was included in the investigation and '(✓)' that the expression/content pair was investigated in Tonhauser et al. 2013. 'CC' abbreviates content of clausal complement. Color coding identifies the class of the content of the English translations: the conventional implicature in orange, presuppositions in magenta, projective non-presupposed content in purple and non-projective content in black.

say that the content of the NRRC is more projective than the CC of *discover* means that the likeliness with which listeners take speakers to be committed to the truth of the content of the NRRC is higher than the likeliness with which listeners take speakers to be committed to the truth of the CC of *discover*.

This paper remains agnostic about the underlying interpretation of projection as a gradient property, though the discussion will be in line with the first interpretation. Across the three investigations in sections 2.1 to 2.3, projection variability was investigated using direct and indirect implication response tasks, which are compatible with either interpretation. To illustrate the two tasks, consider the Paraguayan Guaraní sentence in (4), in which the discontinuous change-of-state expression *nd–...–ve-í-ma* 'not anymore' is embedded in a polar question. The content of interest is the pre-state content, that Marko has smoked.[4]

(4) Context: There is a health program that gives medicine to everybody who has ever smoked or currently smokes. Maria is administering the program in a particular town; since she doesn't know

[4]Paraguayan Guaraní examples are given in the standardized orthography of the language used in Paraguay (Velázquez-Castillo 2004b), except that all postpositions are suffixed to their host. Following this orthography, stressed oral syllables are marked with an acute accent and stressed nasal syllables are marked with a tilde; acute accents are not written for normally accented words, which have stress on the final syllable. I use glosses specified in the Leipzig Glossing Rules and the following additional glosses: A/B = set A/B cross-reference marker, AG = agentive, JE = reflexive/passive marker, NOM.TERM = nominal terminative aspect, NONAG = non-agentive, PE = (in)direct object and locative marker, pron = pronoun, PROSP = prospective aspect.

the people in the town, she is assisted by Mario, a local. Maria hears another person ask Mario:

Márko=pa nd-o-pita-vé-i-ma?
Marko=Q NEG-A3-smoke-more-NEG-PRF

'Does Marko not smoke anymore?' (adapted from Tonhauser et al. 2013:88)

In both direct and indirect implication response tasks, participants are presented with the context and the speaker's utterance. In a direct implication response task, participants are then asked whether, given what the speaker asked, the speaker is certain that Marko has smoked. A 'yes' response would be taken to indicate that the content projected; a 'no' response that it did not. In an indirect implication response task, by contrast, participants are not asked directly about the content of interest, but they are asked a question about the speaker's utterance such that the participant's answer to that question allows the researcher to conclude whether the participant takes the speaker to be certain of the content of interest. For (4), for example, participants can be asked whether Maria, given the question that was put to Mario, would give the medicine to Marko. A positive answer, that, yes, Maria would give the medicine to Marko, would be taken to mean that the speaker was certain that Marko smoked in the past and a negative answer would be taken to mean that the speaker was not certain that Marko smoked in the past.

Direct and indirect implication response tasks can be implemented with binary, categorical response options as well as with gradient ones. In the first case, participants are asked to respond in the affirmative or the negative (e.g., *yes* or *no* in English); this option was used in Tonhauser et al. 2013 and in the investigation in section 2.1. An affirmative response ('yes, the speaker is certain of the relevant content') is taken to indicate that the content projected, a negative one not. With a categorical response task, one content is more projective than another if more participants give affirmative responses to the first than to the second. In the second case, participants are asked to respond on a scale, such as a 5-point Likert scale or a continuous scale; this option was used in Tonhauser et al. 2018 and in the investigations in sections 2.2 and 2.3. Here, the higher the participant's response on the scale, the more projective the content is. With a gradient response task, one content is more projective than another if participants' overall responses are higher for the first than for the second.

2.1 Investigating projection variability in one-on-one elicitation

This section presents the findings of an investigation of projection variability in Paraguayan Guaraní based on categorical responses elicited in direct and indirect implication response tasks in one-on-one meetings with native speakers of the language. A first inkling of projection variability in the language based on categorical responses is already found in Tonhauser et al. 2013. The authors found that the contents associated with 14 expressions were projective, but they also reported variability in the three participants' responses: whereas the three participants agreed for items with 12 of the 14 expression/content pairs, they did not for items with the exclusive *=nte* 'only' and the proximal adverb *haimete* 'almost':[5]

> While one [participant] consistently gave responses on a variety of examples that support the hypothesis that these two [contents] are projective, the other two [participants] gave responses to several examples containing these [expressions] that did not support the hypothesis (in particular when the [expression] was embedded under a modal or occurred in the antecedent of a conditional). Thus, while there is evidence that the prejacent of *=nte* 'only' and the polar [content] of *haimete* 'almost' are projective in Guaraní, we note that their projective behavior may be less robust than that of [contents] of other [expressions] (where the three [participants'] responses strongly agreed with each other). (Tonhauser et al. 2013:89)

[5]This quote has been altered slightly to match the spelling conventions and the terminology used in the current paper.

The findings reported in Tonhauser et al. (2013) suggest that the prejacent of *=nte* 'only' and the polar content of *haimete* 'almost' are less projective than the content associated with the other 12 expressions investigated, which include the contents associated with NRRCs, possessive noun phrases, the change-of-state expression and *–kuaa* 'know'. This section presents the findings of an investigation that, in using the same methods, expands on that of Tonhauser et al. 2013 by investigating the projectivity of the CCs of 14 Paraguayan Guaraní verbs, namely *–kuaa-uka* (know-CAUS) 'point out', *–mo'ã* 'think', *–ha'arõ* 'hope', *–juhu* 'discover', *–topa* 'discover', *–hecha-uka* (see-CAUS) 'show, reveal' and *–hecha-kuaa* (see-know) 'understand, realize', *–hecha* 'see', *–hendu* 'hear', *–vy'a* 'be happy', *–guerotĩ* 'be ashamed', *–ñembyasy* 'regret', *–'e* 'say' and *–mombe'u* 'confess, tell' (see Table 1).

2.1.1 Methods

Participants Four self-reported native speakers of Paraguayan Guaraní participated in the one-on-one elicitation sessions (3 women, 1 man). They were living in San Lorenzo (in the Central department of Paraguay) and were between 20 and about 60 years old. One of these participants also provided responses for the research reported on in Tonhauser et al. 2013.

Materials A total of 18 sets of examples were created with the 14 verbs: a set of examples consisted of an example in which the verb was not embedded under an entailment-canceling operator and variants of that example in which the verb was embedded under one of four entailment-canceling operators, i.e., the Family-of-Sentences variants (Chierchia and McConnell-Ginet 1990). The set for *–hecha* 'see' is given in (5) with the context in which the sentences were presented to the participants. In (5a), *–hecha* 'see' is not embedded under an entailment-canceling operator. The sentences in (5b-e) are the Family-of-Sentences variants of (5a): the verb and its complement are embedded under the possibility modal *i-katu* (B3-possible) 'it's possible' in (5b), in a polar question in (5c), under negation in (5d) and in the antecedent of a conditional in (5d). The example in (5f) is a distractor: each set included one or two examples that did not feature the verb under investigation but were otherwise similar to the examples in the set.

(5) Context: In my house, we only buy milk if there is absolutely no milk left. If there's even just a little bit of milk left, we don't buy any. One day, I come to the house and my sister says to me:

a. Ñande-sy o-hecha ndaipori=ha kamby.
B1pl.incl-mother A3-see not.exist=NMLZ milk
'Our mother saw that there is no milk.'

b. I-katu ñande-sy o-hecha ndaipori=ha kamby.
B3-possible B1pl.incl-mother A3-see not.exist=NMLZ milk
'It is possible that our mother saw that there is no milk.'

c. Ñande-sý=pa o-hecha ndaipori=ha kamby?
B1pl.incl-mother=Q A3-see not.exist=NMLZ milk
'Did our mother see that there is no milk?'

d. Ñande-sy nd-o-hechá-i ndaipori=ha kamby.
B1pl.incl-mother NEG-A3-see-NEG not.exist=NMLZ milk
'Our mother didn't see that there is no milk.'

e. Ñande-sy o-hechá-ramo ndaipori=ha kamby, a-jepy'apý-ta.
B1pl.incl-mother A3-see-if not.exist=NMLZ milk A1sg-upset-PROSP
'If our mother saw that there is no milk, I am going to be upset.'

f. I-katu ñande-sy oi-mo'ã ndaipori=ha kamby.
B3-possible B1pl.incl-mother A3-think not.exist=NMLZ milk
'It's possible that our mother thought that there is no milk.'

The clausal complement was nominalized with *=ha* 'NMLZ', except when the verb was *–ha'arõ* 'hope'.[6] For each of the 14 verbs, one set of examples was created; two exceptions were *–juhu* 'discover' and *–mombe'u* 'confess, tell', for which two and three, respectively, sets of examples were created (as part of an initial attempt to elicit responses on more than one set of examples for each verb). For one of the two sets of examples with *–juhu* 'discover', responses were elicited twice. For *–guerotĩ* 'be ashamed', the set of examples did not include an example in which the verb was embedded in a polar question. There were thus a total of 71 Family-of-Sentences examples with 14 verbs. See Appendix A for the 18 sets of examples.

Procedure For each example of the 18 sets of examples, binary, categorical responses were elicited in direct and indirect implication response tasks from the four participants.[7] For instance, I collected binary, categorical responses in an indirect implication response task for the examples in (5) by asking the four participants whether, given what my sister said, we would buy milk. An affirmative response was taken to indicate that the CC projects and a negative response that it does not. The examples in each set were presented in random order to each participant and interspersed with questions about examples from sets with other verbs and about other topics in the grammar of the language.

2.1.2 Results and discussion

Figure 1 plots the percentage of responses indicating projection of the CCs for each of the 18 sets of examples: the ratios on top of the bars are the number of responses indicating projecting over the total number of responses elicited. For comparison, the figure also includes the results reported in Tonhauser et al. 2013 for main clause content (abbreviated 'MC'), *=nte* 'only', *haimete* 'almost', *–kuaa* 'know', the change-of-state expression and NRRCs; the words on top of these bars indicate whether none, some or all responses indicated projection. For *=nte* 'only' and *haimete* 'almost', the percentage of responses indicating projection was arbitrarily set to 50%. As shown, the CCs of many of the verbs were projective, including the CCs of verbs whose English translations are considered non-factive, namely *–'e* 'say', *–hechauka* 'show, reveal' and *–mombe'u* 'confess, tell'. The three sets for *–juhu* 'discover' and *–mombe'u* 'confess, tell' are distinguished by the appended numbers '*_1*', '*_2*', *_3*'; these number do not reflect order information because, as mentioned above, responses on the 71 examples were elicited in random order from each of the four participants. Only the CC of one set of examples with *–juhu* 'discover' (*juhu_1*) and of the sets of examples with *–ha'arõ* 'hope', *–topa* 'discover' and *–mo'ã* 'think' received 2 or fewer responses indicating projection, suggesting that these CCs were not projective.

In response to question (3a), whether there is projection variability in languages other than English, the findings plotted in Figure 1 are suggestive of projection variability in Paraguayan Guaraní. First, for both presupposed CCs (in purple) and projective non-presupposed CCs (in orange), there are sets of examples for which 100% of responses indicated projection (e.g., *–hecha* 'see' and one set of examples with *mombe'u* 'confess, tell', namely *mombe'u_3*) as well as sets of examples for which less than 75% of responses indicated projection (e.g., *–hechakuaa* 'realize' and *–'e* 'say'). This finding suggests within-class projection

[6]Some of the verbs that can take a clausal complement with *=ha* 'NMLZ' can also take a clausal complement without. What governs the distribution of *=ha* 'NMLZ' is a question for future research.

[7]The native speakers of Paraguayan Guaraní that participated in the work reported in this paper were fluent in Paraguayan Guaraní and Spanish (on bilingualism in Paraguay see, e.g., von Gleich 1993; Fasoli-Wörmann 2002; Stewart 2017). The contexts and response questions were presented in Spanish. For discussions of considerations that go into which language to use in research on languages one does not speak natively see Matthewson 2004 and AnderBois and Henderson 2015.

variability. Second, consider the responses given to the set of examples with *–hecha* 'see', the set of examples with *–hechakuaa* 'understand, realize' and the set *juhu_2* with *–juhu* 'discover'. These sets of examples were identical except for the verb, but, of the 16 responses, 10 indicated projection with *–hechakuaa* 'understand, realize', 11 with *–juhu* 'discover' and 16 with *–hecha* 'see'. This finding suggests that the expression that is associated with the projective content (here, the verb associated with the CC) is a source of projection variability; this is referred to as by-expression projection variability.

Finally, recall that for *–juhu* 'discover' and *–mombe'u* 'confess, tell' responses on more than one set of examples was collected. As shown in Figure 1, of the 16 responses, none indicated projection for one set of examples with *–juhu* 'discover' (*juhu_1*); for the other set of examples, 11 indicated projection one time responses were elicited (*juhu_2*) and 13 indicated projection the other time (*juhu_3*). Of the 16 responses elicited for each of the three sets of examples with *mombe'u* 'confess, tell', 9 indicated projection in one set (*mombe'u_1*), 12 in another set (*mombe'u_2*) and 16 in the third set (*mombe'u_3*). This finding suggests that projection variability in Paraguayan Guaraní is introduced by the lexical content of the complement and/or by the context in which the example is presented. In sum, the findings of this investigation suggest, in response to question (3a) that there is projection variability in Paraguayan Guaraní and, specifically, that there is within-class and by-expression variability as well as by-context and/or by-lexical content variability.

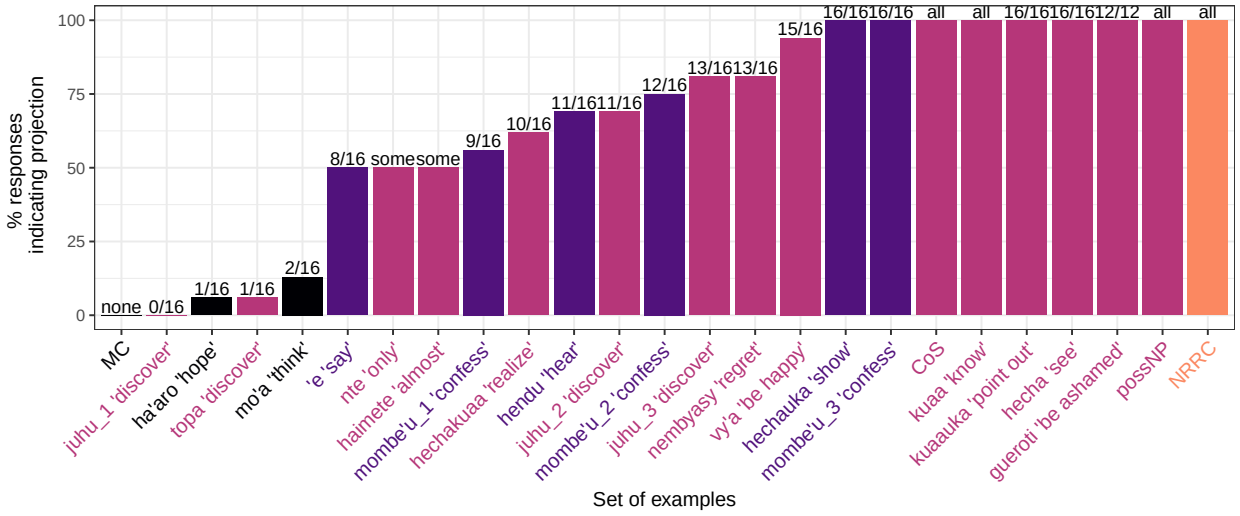

Figure 1: Plot of percentage of responses indicating projection, by set of examples. Ratios on top of bars are the number of responses indicating projection over total number of responses; words on top of bars indicate results reported in Tonhauser et al.'s 2013. 'MC' abbreviates main clause content and 'possNP' posessive noun phrases. Color coding identifies the class of the content of the English translations: the conventional implicature in orange, presuppositions in magenta, projective non-presupposed content in purple and non-projective content in black.

In addition to the aforementioned sources, projection variability was also introduced by the participants; this is referred to as by-participant variability. To illustrate, consider the four participants' responses to the set of examples with *–hecha-kuaa* 'understand, realize', a variant of the set of examples in (5), in Table 2. As shown, the participants vary in their responses to the four examples.

There are limitations, however, in what one can conclude about projection variability on the basis of this investigation. For instance, we cannot comprehensively assess by-expression projection variability because the sets of examples differed not just in the verbs realized but also in the contexts and the complements. Further variation may have been introduced by the fact that both direct and indirect implication response tasks were used. Finally, not enough data were collected to submit the findings to statistical analysis: as a

	Participant			
Entailment-canceling operator	#1	#2	#3	#4
Possibility modal	yes	no	no	no
Polar question	yes	yes	yes	yes
Antecedent of conditional	yes	no	yes	yes
Negation	yes	no	no	yes

Table 2: The four participants' responses to the set of examples with *–hecha-kuaa* 'understand, realize'

consequence, the findings are merely suggestive of projection variability. The investigations presented in the next two sections address these limitations by only using the direct implication response task, by presenting the examples in the same, minimal context and by reducing by-eventuality variability. In the experiment, sufficient data were collected to submit them to statistical analysis.

2.2 Pilot study

This section presents the findings of a pilot study designed to investigate projection variability in Paraguayan Guaraní on the basis of gradient responses in a direct implication response task. In addition to addressing question (3a), the pilot study served to identify whether Paraguayan Guaraní participants are comfortable giving gradient responses. The expressions investigated were embedded in polar questions.

2.2.1 Methods

Participants Five self-reported speakers of Paraguayan Guaraní participated (4 women, 1 man). They were working for me as paid language consultants, living in San Lorenzo (in the Central department of Paraguay) and between 20 and about 60 years old.

Materials Nine polar questions were formed by combining one of the following nine verbs with the same proper name subject and the same complement clause: *–kuaa* 'know', *–guerotĩ* 'be ashamed', *–mombe'u* 'confess, tell', *–vy'a* 'be happy', *–hecha-kuaa* 'understand, realize', *–juhu* 'discover', *–ñembyasy* 'regret', *–mo'ã* 'think' and *–ha'arõ* 'hope'. As shown in (6) for *–kuaa* 'know', the proper name subject was *Sándra* and the lexical content of the complement clause was that Sandra's sister is pregnant. In the polar question with *–ha'arõ* 'hope' the complement was not nominalized with *=ha*.

(6) Sándra=pa oi-kuaa iñ-ermána hy'e guasu=ha?
Sandra=Q A3-know B3-sister stomach big=NMLZ
'Does Sandra know that her sister is pregnant?'

Procedure The pilot study was carried out as part of one-on-one elicitation sessions. Participants were seated in front of me and instructed to imagine that they were overhearing a man called Samuel (who they don't know) talking on his cell phone to somebody who they also don't know. Overhearing people talking on their cell phones was a natural scenario for these participants. Each trial consisted of me reading one of the nine polar questions to the participant. Reading the stimuli to the participants was necessary because most speakers of Paraguayan Guaraní do not have practice reading in the language. For each participant, I began with the polar question with *–vy'a* 'be happy', and proceeded in a random order through the other eight polar questions. In each trial, I asked the participant whether Samuel is certain that Sandra's sister is pregnant, using the Spanish question in (7):

(7) Según lo que dijo Samuel, el esta seguro de que la hermana de Sandra esta embarazada?
according the what said Samuel he is certain that the sister of Sandra is pregnant
'According to what Samuel said, is he certain that Sandra's sister is pregnant?'

Participants were instructed to respond using a 5-point Likert scale, which was presented on a piece of paper. The extreme points of the scale were labeled in Spanish: '1' was labeled with *no esta seguro* 'he's not sure' and '5' was labeled with *esta muy seguro* 'he's very sure'. The 5-point scale was chosen because it was familiar to the participants from the local school system where 1 is the worst grade and 5 the best. The higher a participant's response, the more projective the CC was taken to be.

2.2.2 Results and discussion

Participants did not appear to find it difficult to give responses on the 5-point Likert scale to the response question in (7) and also did not report any difficulties with the task. The pilot study was therefore taken to establish that gradient projectivity ratings can be collected from native speakers of Paraguayan Guaraní. The five participants' responses to the nine polar questions are shown in Table 3, organized in descending order by the by-expression response means. Participant #4 did not judge the polar question with *–ha'arõ* 'hope' to be acceptable (they commented that they took the verb to mean 'wait'). As shown, the response means differ by the verb that was realized in the polar question: it is lowest for *–mo'ã* 'think' and *–ha'arõ* 'hope', and highest for the emotive verbs *–vy'a* 'be happy, *–guerotĩ* 'be ashamed' and *–ñembyasy* 'regret'. In response to question (3a), the findings suggest that there is by-expression projection variability in Paraguayan Guaraní within the class of presuppositions: the response means of the cognitive verbs *–juhu* 'discover' and *-kuaa* 'know' are numerically lower than that of the emotive verbs.

Expression (all verbs)	**Participants' responses** #1	#2	#3	#4	#5	**Response mean**
–vy'a 'be happy'	5	5	5	5	5	5
–guerotĩ 'be ashamed'	5	5	5	5	5	5
–ñembyasy 'regret'	5	5	4	5	5	4.8
–hecha-kuaa 'understand, realize'	5	5	4	5	3	4.4
–mombe'u 'confess, tell'	5	1	5	4	5	4
–kuaa 'know'	2	5	4	5	2	3.6
–juhu 'discover'	3	5	3	5	1	3.4
–mo'ã 'think'	4	3	2	5	2	3.2
–ha'arõ 'hope'	1	1	4	–	1	1.75

Table 3: Participants' responses by verb; '–' means that no response was given. Color coding identifies the class of the content of the English translations: presuppositions in magenta, projective non-presupposed content in purple and non-projective content in black.

Compared to the investigation reported on in section 2.1, the pilot study allowed for a more comprehensive comparison of by-expression projection variability and, indeed, found variability in the projectivity of the CCs of different verbs. This was achieved by only eliciting responses in a direct implication response task, by presenting the utterances in the same, minimal context and by combining the nine verbs with the same subject noun phrase and complement clause. However, variation may have been introduced in the verbal presentation of the polar questions to the participants. A second shortcoming is that insufficient responses were collected to establish whether the numerical by-expression differences are statistically significant. The experiment reported on in the next section addressed these shortcomings and investigated projection for a

broader set of contents. The verb *–ha'arõ* 'hope' was not included because participants might interpret the verb as 'wait'. The experiment also did not include *–guerotĩ* 'be ashamed' and *–ñembyasy* 'regret' because emotive verbs showed little projection variability in the pilot study. The formulation of the Spanish response question was also changed, based on feedback from one of the participants in the pilot study.

2.3 Experiment

The experiment was designed to address question (3a) by investigating the projectivity of the content associated with 17 expressions: NRRCs, possessive noun phrases, *=nte* 'only', *haimete* 'almost', the change-of-state expression and the following 12 verbs: *–kuaa* 'know', *–kuaa-uka* 'point out', *–juhu* 'discover', *–topa* 'discover', *–hecha* 'see', *–hechauka* 'show, reveal', *–hechakuaa* 'understand, realize', *–vy'a* 'be happy', *–hendu* 'hear', *–mo'ã* 'think', *–'e* 'say' and *–mombe'u* 'confess, tell' (see Table 1).

2.3.1 Methods

Participants 30 self-reported native speakers of Paraguayan Guaraní residing in San Lorenzo participated in the experiment (19 female, 11 male). All of the participants also spoke Spanish. They were paid Gs. 20,000 (about $3.60, at the time) for their participation.

Materials Two polar questions each were formed with the 17 expressions, for a total of 34 target polar questions. The polar questions for the CC of *–mo'ã* 'think' and the possession content of possessive noun phrases are given in (8) and (9), respectively. The complete set of polar questions is given in Appendix B.

(8) Target polar questions with *–mo'ã* 'think'

a. Dóra=pa oi-mo'ã Juã o-ñe-monde=ha hína?
Dora=Q A3-think Juan A3-JE-dress=NMLZ PROG
'Does Dora think that Juan is getting dressed?'

b. Davíd=pa oi-mo'ã iñ-ermáno o-jahu=ha hína?
David=Q A3-think B3-brother A3-bathe=NMLZ PROG
'Does David think that his brother is showering/taking a bath?'

(9) Target polar questions with possessive noun phrases

a. Sándra lóro=pa hasy?
Sandra parrot=Q A3.sick
'Is Sandra's parrot sick?'

b. Mário ryguasú=pa o-kuru?
Mario chicken=Q A3-brood
'Is Mario's chicken brooding?'

To minimize variability between the two polar questions for each expression, the lexical contents of the two polar questions were as similar as possible while avoiding repetition of lexical material. For instance, the lexical contents of the complements of *–mo'ã* 'think' in (8) are that Juan is getting dressed and that David's brother is showering, and the lexical contents of the possessive noun phrases in (9) are that Sandra has a parrot and that Mario has a chicken. Furthermore, the 34 lexical contents of the 34 target polar questions had a roughly comparable prior probability, as assessed by a native speaker.[8] Variability was further reduced by

[8] I presented the native speaker with utterances of sentences describing 34 lexical contents and asked them to identify any that seem much more likely or unlikely than others. A couple of lexical contents were replaced based on this speaker's feedback.

only using third person singular proper names and by realizing the 34 target polar questions with the same word order. Specifically, the first expression in each polar question was the third person singular subject; the question clitic *=pa* 'Q' cliticized to this expression, as shown in (8), or to the noun phrase that dominated the proper name, as shown in (9). The only exception were the polar questions with *haimete* 'almost', which were judged in prior elicitation sessions to be more natural with *haimete* 'almost' as the first expression.

In addition to the 34 target polar questions, the experiment included 6 main clause control polar questions, given in Appendix B. These were included to allow participants to use the full response scale as main clause content was hypothesized to be not projective and to assess whether participants understood the task.

The experiment stimuli were recordings of utterances of the 40 target and control polar questions. The utterances were produced by a female talker, a native speaker of Paraguayan Guaraní, who was instructed to produce the polar questions naturally, without special emphasis on any of the expressions of the polar questions. The recordings were made in a quiet room with the built-in microphone of a MacBook Air laptop computer using the Praat software (Boersma and Weenink 2016) with a sampling rate of 44,100Hz.[9]

The 40 stimuli were divided into two blocks, A and B, of 20 stimuli: each block included one of the two stimuli for each of the 17 expression/content pairs and 3 control stimuli. In each block, the 20 target and control stimuli occurred in the same pseudo-randomized order (as shown in Appendix C). Each participant rated all 40 stimuli: half of the participants were first presented with the stimuli in block A and then the stimuli in block B; the other half of the participants were first presented with the stimuli in block B in reverse order and then with the stimuli in block A in reverse order.

To assess projection, participants were asked a question, in Spanish, about the relevant content of each of the 40 polar question stimuli. The response question for the target stimulus in (8a) is shown in (10):

(10) Según lo que preguntó Magda, que certeza tiene de que Juan se esta vistiendo?
according.to what asked Magda what certainty has.3 of that Juan SELF is dress.PART
'According to what Madgda asked, how certain is she that Juan is getting dressed?' (lit.: '...what certainty does she have that Juan is getting dressed?')

I recorded myself asking the response questions under the aforementioned recording conditions. For each polar question stimulus, a composite recording was created by splicing the recording of the polar question with a one-second silence, and then with the recording of the corresponding response question.

To familiarize participants with the task, three additional polar questions were recorded with the same female talker as well as corresponding response questions. These familiarization stimuli were presented in the same order to all participants.

Procedure Participants were tested individually in a quiet location and were randomly assigned to one of the two block orders. They were seated across from me and shown a picture of a woman talking on a cell phone. They were told that the woman is called Magda and asked to imagine that they are overhearing Magda asking questions of somebody that they don't know on her cell phone. After familiarizing the participant with the task using the three familiarization stimuli, the 40 stimuli were presented one at a time using a Macintosh laptop. The majority of participants agreed to wear Apple EarPods. They were permitted to listen to each stimulus as often as they wanted before responding, but usually listened to stimuli only once.

Participants gave their responses on a 5-point Likert scale which was presented to them on a piece of paper. The scale was labeled in Spanish at three points: '1' with *ninguna certeza* 'no certainty', '3' with *poca certeza* 'low certainty' and '5' with *mucha certeza* 'a lot of certainty'. The higher a participant's response to a polar question stimulus, the more projective the content. Participants gave their responses

[9]These recordings as well as the data and the R code for generating the figures and analyses presented in this paper are available at [redacted].

verbally in Spanish or manually, either by pointing to a number on the Likert scale or by holding up between one and five fingers. They were offered a break after every ten trials.

Data exclusion Prior to analysis, the responses to the one of the target polar question stimuli with *–vy'a* 'be happy' were excluded because the CC was revealed in the course of the experiment to be an idiom and ambiguous between the meanings 'is not hungry' and 'is not stingy'. Furthermore, one participant judged the response questions of two target polar questions to not be targeting content conveyed by the polar questions, and hence did not give a response to those two items. Finally, the responses by one participant were excluded because they responded '5' in all 40 trials, suggesting that they did not understand the task or interpreted the task differently. The analysis is based on the data from the remaining 29 participants.

2.3.2 Results

The mean certainty ratings for the 17 expressions are plotted in Figure 2 with bootstrapped 95% confidence intervals, collapsing over the two polar questions per expression. The jittered gray dots indicate the participants' responses. As expected, the mean certainty ratings for the main clause controls and the CC of *–mo'ã* 'think' are the lowest, at 2.29 and 2.55, respectively. The observation that the mean certainty ratings for the remaining expressions range from 3.38 for *–hecha* 'see' to 4.47 for possessive noun phrases is suggestive of by-expression projection variability as well as projection variability within the class of presupposed contents (in purple). Finally, the projectivity of the conventional implicature of the NRRC (mean rating of 4.38; in light blue) appears to be higher than at least some of the presuppositions.

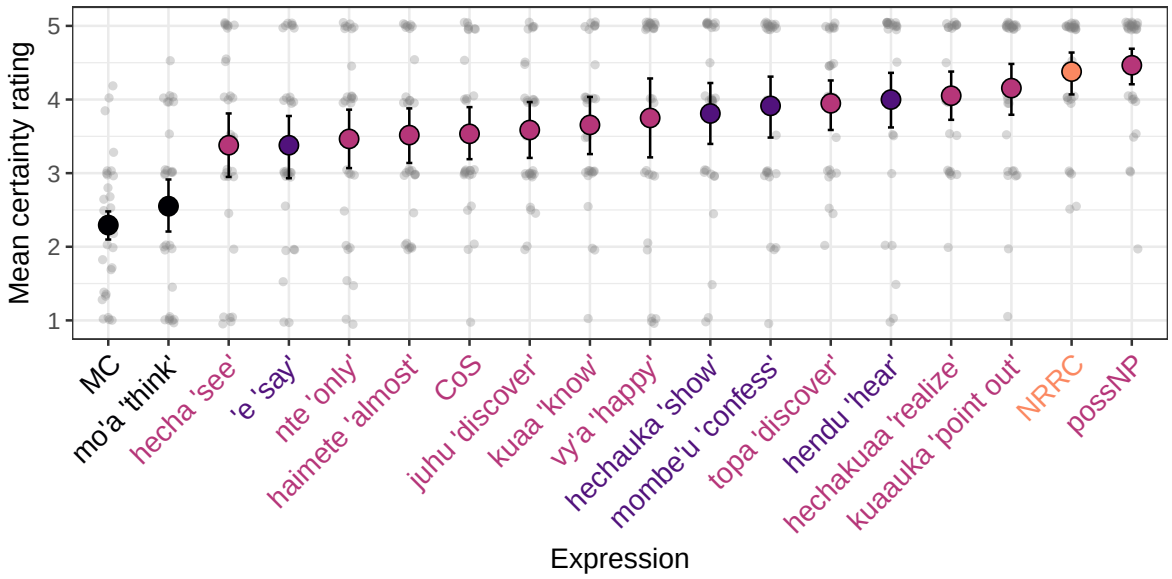

Figure 2: Mean certainty ratings by expression, including main clauses, with 95% bootstrapped confidence intervals, collapsing over lexical content. Jittered gray dots indicate participants' responses. 'MC' abbreviates main clause, 'CoS' change-of-state expression and 'possNP' possessive noun phrase. Color coding identifies the class of the content of the English translation: the conventional implicature in orange, presuppositions in magenta, projective non-presupposed content in purple and non-projective content in black.

Projection variability was analyzed by fitting a mixed-effects ordinal regression model predicting participants' responses from a fixed effect of expression, with main clause controls as the reference level (1,129 data points), using the `clmm` package (Christensen 2013) in R (R Core Team 2016, version 3.5.0, 2018-04-

23). The model included the maximal random effects structure that allowed the model to converge: random by-participant intercepts (capturing participants' variability in responses) and random by-item intercepts (capturing variability in responses by item).[10] Pairwise differences between expressions ($\alpha = .05$) were identified by the `lsmeans` package (Hothorn et al., 2008) using the Tukey multiple-comparison test.

The analysis revealed significant differences between the main clause controls and all other expression/content pairs ($p < .001$ for all), except *–mo'ã* 'think' ($p = 1$). Under the assumption that main clause content is not projective, this finding suggests that the CC of *–mo'ã* 'think' is not projective, as expected, but that all other contents investigated are at least mildly projective. This includes the Paraguayan Guaraní verbs whose English translations are considered non-factive, but whose CCs have been observed to be projective in, for instance Schlenker 2010; Anand and Hacquard 2014; Spector and Egré 2015 and Tonhauser et al. 2018, namely *–e'* 'say', *–hechauka* 'show, reveal', *–mombe'u* 'confess, tell' and *–hendu* 'hear'. The analysis furthermore revealed significant differences between possessive noun phrases and *–hecha* 'see' ($p = .003$), *–'e* 'say' ($p = .003$), *=nte* 'only' ($p = .01$), *haimete* 'almost' ($p = .01$), the change-of-state expression ($p = .01$) and *–juhu* 'discover' ($p = .03$); a marginal difference was observed with *–kuaa* 'know' ($p = .08$). This finding suggests that there is projection variability within the class of presuppositions in Paraguayan Guaraní: the relevant content of possessive noun phrases is more projective than the relevant contents of the change-of-state expression, *haimete* 'almost', *–hecha* 'see', *–juhu* 'discover' and *=nte* 'only'. Finally, the analysis revealed significant differences between NRRCs and *–hecha* 'see' ($p = .01$), *–'e* 'say' ($p = .01$), *=nte* 'only' ($p = .037$), *haimete* 'almost' ($p = .047$) and the change-of-state expression ($p = .047$); a marginal difference was observed with *–juhu* 'discover' ($p = .09$). This finding suggests that there is between-class projection variability: the content of NRRCs, a conventional implicature, is more projective than some presuppositions. In sum, in response to question (3a), there is projection variability in Paraguayan Guaraní both between and within classes of projective content.

Likelihood ratio tests furthermore revealed that by-participant intercepts were justified (sd = .67, $p < .0001, \chi^2(1)$) but by-item intercepts were not (sd = 0, $p = 0.99, \chi^2(1)$). That is, projection variability was introduced by participants but not by lexical contents. Figure 3, which plots the mean certainty ratings of each of the four classes of content by participant, illustrates the observed by-participant projection variability.

2.4 Summary and discussion

The three investigations have established that utterance content in Paraguayan Guaraní varies in its projectivity. One of the sources of projection variability is the expression associated with the content:

(11) Summary of by-expression projection variability in Paraguayan Guaraní

a. One-on-one elicitation (section 2.1)
CCs of verbs: The CCs of the 14 verbs included ones that are highly projective and ones that are less highly projective, both among the verbs whose English translations are factive and among the verbs whose English translations are non-factive.

b. Pilot study (section 2.2)
CCs of verbs: The CCs of the 9 verbs differed, at least numerically, in how projective they are, with the CCs of emotive verbs highly projective and those of cognitive verbs less projective.

c. Experiment (section 2.3)
– **CCs of verbs:** The CC of *–mo'ã* 'think' is not projective, whereas the CCs of the other verbs investigated are projective and differ, at least numerically, in how projective they are. This

[10] A model with random slopes for expression by participant did not converge. A model that included the interaction of expression and block order was not significantly better than reported model, suggesting that the order in which participants completed the task did not influence their responses.

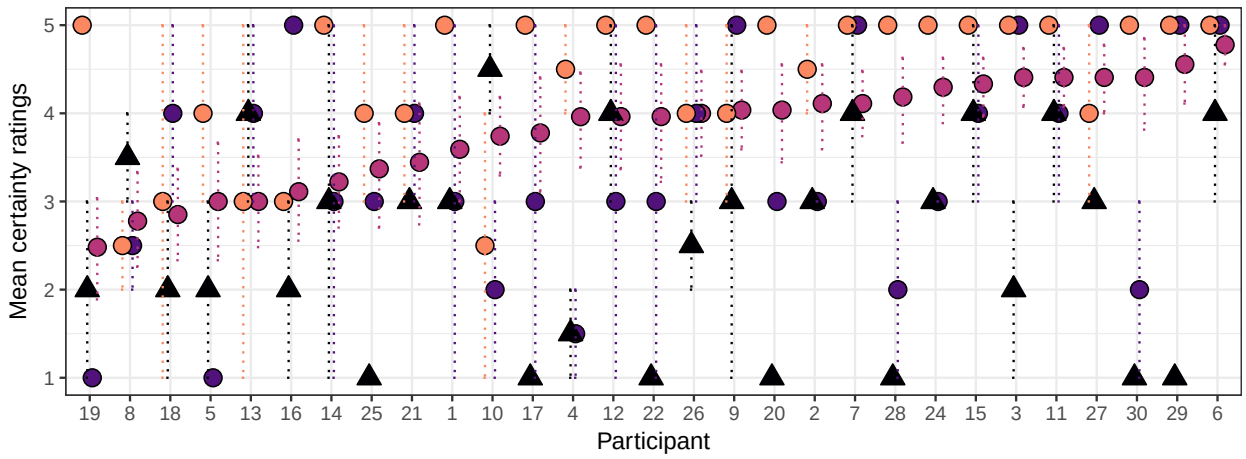

Figure 3: Mean certainty ratings of the four classes of content by participant, excluding main clauses, with 95% bootstrapped confidence intervals given as dotted lines. Participants ordered left-to-right by increasing mean response to presupposition items. Color and shape coding identifies the class of the content of the English translations: the conventional implicature in orange circles, presuppositions in magenta circles, projective non-presupposed content in purple circles and the non-projective content in black triangles.

includes the CCs of verbs whose English translations are non-factive but projective, namely *–'e* 'say', *–mombe'u* 'confess, tell', *–hechauka* 'show, reveal' and *–hendu* 'hear'.

- **Presuppositions:** There is projection variability among contents typically considered presuppositions: the possession content of possessive noun phrases is more projective than the pre-state content of the change-of-state expression, the polar content of *haimete* 'almost', the prejacent of *=nte* 'only' and the CCs of *–hecha* 'see' and *–juhu* 'discover'.
- **Conventional implicature vs. presuppositions:** The conventional implicature associated with NRRCs is more projective than several presuppositions, including the CCs of *–hecha* 'see' *–'e* and 'say', the prejacent of *=nte* 'only', the polar content of *haimete* 'almost' and the pre-state content of the change-of-state expression.

The three investigations also established that projection variability is contributed by participants varying in their responses (by-participant projection variability) and may also be introduced by the context and the lexical contents (by-context and by-lexical content projection variability). In sum, in response to question (3a), we conclude that projection variability is attested in Paraguayan Guaraní and is not a property unique to English and German. Given that Paraguayan Guaraní, as a Tupí-Guaraní language, is genetically unrelated to and typologically different from the Germanic languages, it would be surprising if projection variability just happened to be observed in these three languages, rather than being a more widespread phenomenon. We can tentatively hypothesize that projection variability is a cross-linguistically universal property of projective content, though of course investigations in many other languages are needed to verify this hypothesis.

Assessing projection analyses We can assess the empirical adequacy of current analyses of different classes of projective content against the projection variability observed in Paraguayan Guaraní. For conventional implicatures, analyses generally predict that such content is highly projective because it is independent (and thereby does not end up in the semantic scope of entailment-canceling operators) or directly updates the common ground (e.g., Potts 2005; Murray 2014; AnderBois et al. 2015; Gutzmann 2015). The cross-linguistic applicability of these analyses is supported by the finding that the content of NRRCs in Paraguayan

Guaraní is highly projective.

For presuppositions, some contemporary research distinguishes between 'hard' and 'soft' presupposition triggers (see also Kadmon 2001), though there is no consensus on which triggers are hard and which are soft (as discussed in Tonhauser et al. 2018). The presuppositions of hard triggers are required to be entailed by or satisfied in the common ground of the interlocutors (Heim 1983; van der Sandt 1992; Beaver and Krahmer 2001; though see Abrusán 2016). This lexically specified requirement correctly predicts that the possession content of possessive noun phrases in Paraguayan Guaraní is highly projective. The presuppositions of soft triggers, on the other hand, are not lexically specified but rather derived through lexical specification of alternatives and pragmatic mechanisms (e.g., Abusch 2002, 2010; Chemla 2009; Romoli 2015) or through pragmatic mechanisms alone (e.g., Simons 2001; Abrusán 2011, 2016; Simons et al. 2017). The English translations of at least three of the Paraguayan Guaraní expressions whose associated contents are less projective have been assumed to be soft triggers, namely the pre-state content of the change-of-state expression (English *stop*) and the CCs of the verbs *–juhu* 'discover' and *–hecha* 'see'. Thus, analyses that distinguish hard and soft triggers can be assumed to account for some of the projection variability observed among presuppositions in Paraguayan Guaraní. By contrast, analyses that do not distinguish between soft and hard triggers cannot capture the observed variation. It is not clear, however, whether the distinction between soft and hard triggers suffices to capture the projection variability observed: the English translations of *–topa* 'discover' and *–hechakuaa* 'realize' are also considered soft triggers, but their CCs were not consistently less projective than the possession implication of possessive noun phrases. Future research needs to investigate which classes of projective content are distinguished by projection variability in Paraguayan Guaraní.

Finally, consider the verbs *–'e* 'say', *–mombe'u* 'confess, tell', *–hechauka* 'show, reveal' and *–hendu* 'hear'. As mentioned, their English translations are taken to be non-factive and some researchers have argued that the CCs of the English verbs are projective. We now have experimental evidence that the CCs of the Paraguayan Guaraní translations are projective (and from Tonhauser et al. 2018 for *confess*). This finding provides impetus to include such verbs in investigations of projective content cross-linguistically. To capture the projectivity of the CCs of such verbs in English, Schlenker (2010) and Spector and Egré (2015) proposed that these verbs are ambiguous between a factive lexical entry, which predicts projection, and a non-factive one, which does not. While this analysis would capture that the CCs of these verbs is projective, it does not capture that the CCs of these verbs are just as projective as the CCs of verbs that are taken to only have a factive lexical entry, like *–kuaa* 'know' or *–vy'a* 'be happy'. A more promising avenue may be to derive the projectivity of the CCs of such verbs from their discourse uses, as suggested by Anand and Hacquard (2014). Developing a fully-fledged analysis along these lines is a topic for future research.

In sum, currently available analyses can predict at least some of the projection variability observed in Paraguayan Guaraní, like the distinction between conventional implicatures and (some) presuppositions, as well as some differences among presuppositions. But because investigations of projection variability are still in their infancy, not just in Paraguayan Guaraní but also in well-studied languages like English, further investigations are needed to fully understand which classes of projective content are distinguished by projection and how to capture the projection of contents that have largely been neglected up to now.

Methodological implications Projection variability in Paraguayan Guaraní was explored on the basis of three investigations whose methods differed in a number of ways, as summarized in Table 4. The data collected through one-on-one elicitation were the most heterogeneous: both direct and indirect implication responses were collected, four entailment-canceling operators were used and context, prosody and lexical content were not controlled for (as different items may have had different contexts, different participants may have been presented with different prosodies, and the lexical content may have had different prior probabilities across the items). As is typical for this type of investigation, classically known as 'fieldwork',

the data were not subjected to statistical analysis. By comparison, the data collected in the pilot study and the experiment were more homogenous; the latter were subjected to statistical analysis.

	One-on-one elicitation (section 2.1)	**Pilot study** (section 2.2)	**Experiment** (section 2.3)
Implication response task	direct and indirect	direct	direct
Response options	yes/no	5-point Likert scale	5-point Likert scale
Entailment-canceling operator	negation, polar question, modal, antecedent of conditional	polar question	polar question
Number of expression/content pairs	14	9	18
Items per expression/content	4-12	1	2 (6 for main clauses)
Participants	4	5	30
Context controlled for?	no	yes	yes
Prosody controlled for?	no	no	yes
Lexical content controlled for?	no	yes	yes
Data statistically analyzed?	no	no	yes

Table 4: Comparison of the methods of the three investigations of projection variability.

Regardless of whether projection variability is investigated in a well-studied language like English or a lesser-studied language like Paraguayan Guaraní, researchers designing such an investigation must strike a balance between several needs. A first need is to engage with the participants in a culturally appropriate manner: in some cases, this may preclude the use of written stimuli or the presentation of stimuli via a computer or head phones; in others, it may mean that binary response options are more natural than gradient ones. A second need is to present participants with tasks that they are comfortable with and can complete with confidence. This may, for instance, preclude the use of direct implication response tasks. Finally, there is the need of the researcher to control for factors that influence projection: the more factors are controlled for, the easier it is to pinpoint observed variability to a specific factor. The three investigations differed in the extent to which these needs were met. In the investigation based on one-on-one elicitation, the items were heterogeneous and different response tasks were used to keep the participants engaged, but this heterogeneity limited the ability to pinpoint projection variability to a specific factor. It was easier to pinpoint the observed projection variability in the pilot study, but responses on only 9 items were collected because the task was more repetitive for the participants. Finally, in the experiment, the items were more heterogeneous again in their lexical contents to make the task less repetitive for the participants.

This comparison of the methods used illustrates three of the many options that researchers have available to them in designing an investigation of projection variability. In particular, it is possible to design an investigation based on one-on-one elicitation that controls for more factors than the investigation presented in section 2.1. One can, for instance, present participants in one-on-one elicitation with items that are tightly controlled for in the entailment-canceling operator, the context, the prosody and the lexical content, without making the task repetitive for them, by spreading these items out across several elicitation sessions and interspersing questions about items that concern other empirical phenomena. If sufficient data are collected (even from a small set of participants), such data could even be submitted to statistical analysis. This reinforces the point that so-called fieldwork-based data can be as tightly controlled for as data collected in experiments, even when the investigation is not conducted as a classical experiment.

3 Comparing projection variability in Paraguayan Guaraní and English

Given that we now know that there is projection variability in Paraguayan Guaraní, we can start addressing question (3b) through a comparison of Paraguayan Guaraní and English:

(3a) If projection variability is cross-linguistically attested, is projective content associated with a translation pair similarly projective or are there systematic cross-linguistic differences?

As mentioned in section 1, comparing the two languages is made possible by the methodological similarities between the pilot study (section 2.2), the experiment (section 2.3) and Tonhauser et al.'s 2018 experiments. Specifically, participants in Tonhauser et al.'s Experiments 1a and 1b were presented with polar questions in a minimal context and gave a direct implication response on a gradient scale, like the Paraguayan Guaraní participants. In both sets of investigations, higher responses indicated higher projectivity, though the English participants responded on a sliding scale (coded 0-1) instead of a 5-point Likert scale.

This section compares by-expression projection variability in the two languages by comparing the relative projectivity of the pairs of Paraguayan Guaraní and English expression/content pairs in (12).[11] The pairs in 1.-8. and 12.-13. are translation pairs; for the pairs in 9.-11. the translations are reasonably similar: *–topa* 'discover' is compared to English *find out* so as to not compare both *–juhu* 'discover' and *–topa* 'discover' to the same English verb, and the emotive verbs *–vy'a* 'be happy' and *–ñembyasy* 'regret' are compared to the emotive verbs *be amused* and *be annoyed*, respectively, as they also convey positive and negative emotions, respectively. Finally, contrary to *–hecha-uka* (*see*-CAUS) 'show, reveal' (12.), which was classified as a non-factive verb whose CC can project (like English *show*, coded in purple), *reveal* is typically considered a factive verb (coded in magenta).

(12)
1. Content of NRRCs
2. Possession content of possessive noun phrases
3. Prejacent of exclusive clitic *=nte* 'only' / *only*
4. Pre-state content of the change-of-state expression / *stop*
5. CC of *–kuaa* 'know' / *know*
6. CC of *–hecha* 'see' / *see*
7. CC of *–hecha-kuaa* (see-know) 'understand, realize' / *realize*
8. CC of *–juhu* 'discover' / *discover*
9. CC of *–topa* 'discover' / *find out*
10. CC of *–vy'a* 'be happy' / *be amused*
11. CC of *–ñembyasy* 'regret' / *be annoyed*
12. CC of *–hecha-uka* (see-CAUS) 'show, reveal' / CC of *reveal*
13. CC of *–mombe'u* 'confess, tell' / *confess*

Figure 4 plots the projection means of the English expression/content pairs against the projection means of the Paraguayan Guaraní expression/content pairs, by translation pair. The means of 18 expression/content pairs are plotted because ratings for the CCs of the five verbs *–juhu* 'discover', *–hechakuaa* 'realize', *–kuaa* 'know', *–mombe'u* 'confess, tell' and *–vy'a* 'be amused' were collected in both the pilot study and the

[11] Projection variability can also be compared by investigating which expression/content pairs are statistically significantly different from one another across two languages. This avenue is not pursued here because the data collected in the pilot study in section 2.2 were not subjected to statistical analysis and because of the difference in power between the experiment in section 2.3 and the experiments in Tonhauser et al. 2018. Preliminary evidence for similarities in projection variability comes, however, from the observation that the contents associated with NRRCs and possessive noun phrases in both languages were more projective than the contents associated with the change-of-state expression/*stop*, *=nte* 'only'/*only* and *–juhu* 'discover' / *discover* (though the latter was marginal in Paraguayan Guaraní for NRRCs).

experiment; the Paraguayan Guaraní expressions are marked with '_p' (for pilot) and '_e' (for experiment), respectively. Overall, there is a positive relationship between projection means in the two languages: if one expression/content pair had a numerically higher projection mean than another, the translation of the first expression/content pair also had a numerically higher projection mean than the translation of the second. For instance, the CC of *ñembyasy*/*be annoyed* is numerically more projective than the possession content of possessive noun phrases, which in turn is numerically more projective than the CC of *kuaa*/*know*, which in turn is numerically more projective than the CC of *hecha*/*see*. A positive relationship is not observed for all translation pairs: for instance, the prejacent of *only* is numerically more projective than the CC of *confess*, but the prejacent of *=nte* 'only' is numerically less projective than the CC of *–mombe'u* 'confess, tell'.[12] These qualitative observations were borne out by Spearman's correlations that were run on the 13 pairs of projection means: there was a moderate, positive correlation between the projection means in the two languages, regardless of whether the projection means for the five aforementioned Paraguayan Guaraní verbs came from the pilot (r_S = .556, n = 13, p = .048) or the experiment (r_S = .584, n = 13, p = .036; Cohen 1988). This finding suggests that there are similarities in the by-expression projection variability in Paraguayan Guaraní and in English.

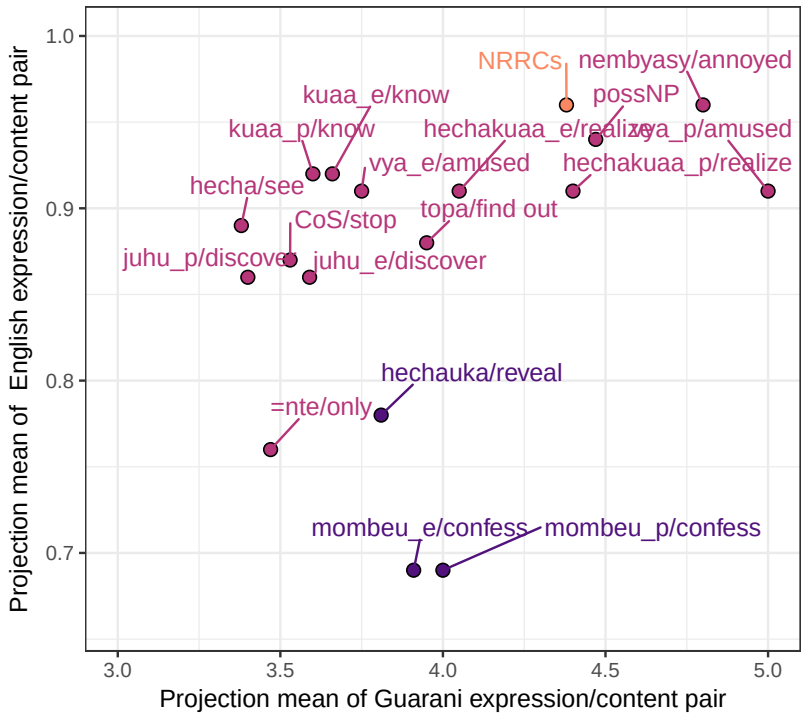

Figure 4: Projection means of English expression/content pairs against projection means of Paraguayan Guaraní expression/content pairs, by translation pair. 'possNP' abbreviates possessive noun phrases and 'CoS' abbreviates change-of-state expression. Color coding identifies the class of the content of the Paraguayan Guaraní expressions: the conventional implicature in orange, presuppositions in magenta and projective non-presupposed content in purple.

[12]Future research needs to determine what such differences are due to. One possibility is that such differences are due to the lexical contents of the items used in the experiment and in Tonhauser et al.'s 2018. Another possibility is that the truth conditional content and/or discourse use of *–mombe'u* 'confess, tell' and *confess* are different.

The finding that projection variability in Paraguayan Guaraní and English is similar has strong implications for analyses of projective content because only some lead us to expect such similarities. Consider first analyses of presuppositions.[13] On lexicalist analyses, that a content is a presupposition is lexically specified for the expression associated with the content, the so-called presupposition 'trigger' (e.g., Heim 1983; van der Sandt 1992; Beaver and Krahmer 2001). For instance, both *–juhu* 'discover' and *discover* would be analyzed as triggers, which means that the lexical entries of these verbs specify that the CC is presupposed. These analyses thereby predict the projectivity of the CCs, but they do not lead us to expect that the CCs are projective in both languages: if the CC of *discover* was projective but the CC of *–juhu* 'discover' was not, that would be just as easy to capture, namely by only specifying the CC of *discover* as a presupposition in the lexical entry of the verb. Thus, lexicalist analyses neither capture projection variability (see section 2.4) nor lead us to expect the observed similarities between the two languages. The same is true of analyses that derive presuppositions from lexical specifications of alternatives and pragmatic mechanisms (e.g., Abusch 2002, 2010; Chemla 2009; Romoli 2015). Such analyses can capture that the relevant contents in both languages are projective but they do not lead us to expect the observed similarities: it would be just as easy to capture that a content is projective when associated with an expression in one language but not with the translation in the other.

By contrast, analyses of presuppositions like those in Abrusán 2011, 2016 and Simons et al. 2017 do lead us to expect cross-linguistic similarities because presuppositions are derived from the interaction of truth-conditional meaning and discourse. Under Abrusán's (2011, 2016) analysis, a content is presupposed if it is a lexical entailment as well as backgrounded in the context in which the expression associated with the context is uttered. In Simons et al. 2017, the projection of a content is derived from it being entailed by the question under discussion addressed by the utterance of the sentence that realizes the expression associated with the content. Thus, under these analyses, it would be surprising to find that the CC of one verb is entailed and projective but the CC of the other is entailed but not projective: if the CCs of both *discover* and *–juhu* 'discover' are entailed content, we also expect both of them to be projective. Additionally, such analyses also lead us to expect cross-linguistic similarities in how projective a content is: if *discover* and *–juhu* 'discover' are similar in their discourse uses, the analyses predict that their CCs are similarly projective.

Next we consider NRRCs: as shown in Figure 4, the content of NRRCs is highly projective in both Paraguayan Guaraní and English. One of the defining properties of the content of English NRRCs is that it is independent (Potts 2005), which means that it does not contribute to the calculation of the truth conditions of the clause in which the NRRC occurs. That NRRCs in Paraguayan Guaraní also have this property was shown in Tonhauser et al.'s 2013 discussion of obligatory local effect. By contrast, presuppositions like the CCs of *–juhu* 'discover' and *discover* are not independent and have obligatory local effect in both languages. Formal analyses capture this property of the content of NRRCs by assuming that it is contributed to a separate dimension of meaning or directly updates the common ground without otherwise participating in the compositional derivation of the content of the clause (e.g., Potts 2005; Murray 2014; AnderBois et al. 2015; Gutzmann 2015). That the content of NRRCs in both Paraguayan Guaraní and English is highly projective falls out of either analysis as a consequence of the content being independent, as a consequence of which it does not end up in the semantic scope of entailment-canceling operators. It would be surprising to find that the content of NRRCs in one language was independent and projective, but independent and not projective in another language, thereby leading us to expect the observed cross-linguistic similarities.

In sum, analyses of projective content differ in the extent to which they lead us to expect the observed similarities in projection variability in Paraguayan Guaraní and English. In general, analyses on which projection is derived from other properties of the projective content, such as entailment for presuppositions or independence for conventional implicatures, lead us to expect the similarities, whereas analyses that

[13]The correlation also holds for the 11 expression/content pairs coded as presuppositions, again regardless of whether the projection means come from the pilot (r_S = .633, n = 11, p = .036) or the experiment (r_S = .624, n = 11, p = .04).

derive projection from lexical specification do not.

The observed similarities in projection variability provide further empirical support for the claim that at least some projective content is nondetachable (see, e.g., Levinson and Annamalai 1992; Simons 2001; Tonhauser et al. 2013). Following Grice's 1975 terminology, what it means for projectivity to be nondetachable is that the projectivity of the content is not due to the expression that is associated with the content but rather due to the truth-conditional meaning of the expression in combination with its discourse use. To date, two types of empirical support for the nondetachability of projectivity have been provided. First, there is within-language evidence. As Simons (2001) pointed out, the pre-state content of the sentences in (13), that Juana has been laughing, is projective regardless of which of the English change-of-state expressions is used. Consequently, the argument goes, it is more parsimonious to assume that the projectivity of this content is due to the truth-conditional meaning shared by *stop, quit* and *cease* than to assume that projectivity is due to the expressions themselves. The second type of evidence is cross-linguistic. Tonhauser et al. (2013) showed, for instance, that the pre-state content of the Paraguayan Guaraní translation of (13) in (14) is projective: this is expected if the projectivity of the pre-state content is not due to the change-of-state expression but rather due to the truth-conditional meaning shared by the Paraguayan Guaraní and English expressions. In Levinson and Annamalai 1992, this point is made based on Tamil (Dravidian) translations of English expressions.

(13) Did Juana { stop / quit / cease } laughing? (adapted from Simons 2001:435)

(14) Juána=pa nd-o-puka-vé-i-ma?
Juana=Q NEG-A3-laugh-more-NEG-PRF
'Is Juana not laughing anymore?'

This paper provides a third type of empirical support for the nondetachability of projection: the projective content associated with translation pairs is similar in how projective it is. For instance, not only is the pre-state content of two change-of-state expressions projective within a language and across two languages, but it is also similar in how projective it is across two languages. Thus, the observation that projection variability is similar in Paraguayan Guaraní and English provides further impetus to developing analyses of projective content that derive projection from the truth-conditional meaning of expressions and their discourse use, so as to predict the nondetachability of projective content.

4 Conclusions

This paper presented the findings of three investigations of projection variability in Paraguayan Guaraní, with the goal of starting to fill the lacuna of cross-linguistic research on projection variability:

(3) Questions about projection variability across languages
   a. What is the typological profile of projection variability? Is it unique to English, a cross-linguistically universal property of projective content, or observed in some but not all languages?
   b. If projection variability is cross-linguistically attested, is projective content associated with a translation pair similarly projective or are there systematic cross-linguistic differences?

The three investigations showed, in response to question (3a), that there is projection variability in Paraguayan Guaraní: specifically, it was established that there is by-expression and by-participant projection variability, as well as variability introduced by context and/or lexical content. The comparison of Paraguayan Guaraní and English showed, in response to question (3b), that there is a positive relationship between the projectivity of content in the two languages: the more projective that content associated with an expression is, the more projective it also is when associated with the translation of the expression in the other language.

Taken together, both findings contribute to our understanding of what empirically adequate analyses of different classes of projective content look like. In the future, analyses will have to be assessed against whether they can account for projective content in Paraguayan Guaraní and the observed similarities between Paraguayan Guaraní and English. Of course, investigations of projection variability in many more diverse languages are needed to understand the typological profile of projection variability and whether it is a cross-linguistically universal property of projective content. It is hoped that the methodological insights gained in the three investigations of Paraguayan Guaraní facilitate research on projective content in other languages, whether in the field or in the laboratory.

A Examples used in one-on-one elicitation

This appendix provides the unembedded sentences from which the Family-of-Sentences variants were formed for the investigation described in section 2.1 as well as the English translations of the response questions in the direct and indirect implication response tasks that the participants were asked. Only the examples not yet discussed in section 2.1 are given.

(15) Context: Growing up, my friend Tina never knew whether she wanted to have kids. We lost touch over the last years but I know that she recently got married. I don't know if she and her husband have decided to have kids or not, but I recently ran into Tina's mother and she told me:

a. **–vy'a 'be happy'**

Tína o-vy'a ha'e hy'e guasu=ha.
Tina A3-happy pron.3.AG stomach big=NMLZ

'Tina is happy that she is pregnant.'
Response question: Given what Tina's mother says, is she affirming that Tina is pregnant? (15/16 positive responses to Family-of-Sentences variants)

b. **–topa 'discover'**

Tína o-topa ha'e hy'e guasu=ha.
Tina A3-discover pron.3.AG stomach big=NMLZ

'Tina discovered that she is pregnant.'
Response question: Given what Tina's mother says, is she affirming that Tina is pregnant? (1/16 positive responses to Family-of-Sentences variants)

(16) **–hendu 'hear'**
Context: Maria is a fan of a particular soccer team. The team played yesterday. I run into Maria's husband, who tells me:

María o-hendu ij-ekípo o-perde=ha-gue.
Maria A3-hear B3-team A3-lost=NMLZ-NOM.TERM

'Maria heard that her team lost.'
Response question: Given what Maria's husband says, is he affirming that Maria's team lost? (11/16 positive responses to Family-of-Sentences variants)

(17) **–ñembyasy 'regret'**
Context: In our town, everybody who goes to church has to donate. Raul is visiting our town. I say:

Raul o-ñembyasy o-ho=ha-gue tupa'ó-pe.
Raul A3-regret A3-go=NMLZ-NOM.TERM church-PE

'Raul regrets that he went to church.'
Response question: Given what I say, does Raul have to donate? (13/16 positive responses to Family-of-Sentences variants)

(18) **–mo'ã 'think'**
Context: Marko is in Asuncion. He is waiting on the street for his bus to take him home. There are two women who are waiting with him. One says to the other:

Oi-mo'ã hasẽ=ha-gue Júlia.
A3-think B3.cry=NMLZ-NOM.TERM Julia

'S/he thinks that Julia was crying.'
Response question: Given what the woman says, is she sure that Julia was crying? (2/16 positive responses to Family-of-Sentences variants)

(19) **–guerotĩ 'be ashamed'**
Context: In our school, students have to be quiet from 2 to 3 and focus on their homework. I supervise quiet time but I have to leave briefly so I leave Luli in charge. We have to punish the students who speak during quiet time. When I come back, I ask Luli how it went. She says:

Samuel o-guerotĩ o-ñe'ẽ=ha-gue.
Samuel A3-ashamed A3-speak=NMLZ-NOM.TERM

'Samuel is ashamed that he spoke.'
Response question: Given what Luli says, do I have to punish Samuel? (12/12 positive responses to Family-of-Sentences variants)

(20) Context: My daughter Laura has a boyfriend. I worry about her and her boyfriend, but she doesn't tell me anything about what's going on anymore. I hear my other daughter talk on the phone. She says:

a. **–kuaa 'know'**

Láura oi-kuaa-uka i-kichihá-me hy'e guasu=ha.
Laura A3-know-CAUS B3-boyfriend-PE stomach big=NMLZ

'Laura pointed out to her boyfriend that she is pregnant.'
Response question: Given what my daughter says about her sister, is she sure that her sister is pregnant? (16/16 positive responses to Family-of-Sentences variants)

b. **–hechauka 'show, reveal'**

Láura o-hecha-uka i-kichihá-me hy'e guasu=ha.
Laura A3-see-CAUS B3-boyfriend-PE stomach big=NMLZ

'Laura revealed to her boyfriend that she is pregnant.'
Response question: Given what my daughter says about her sister, is she sure that her sister is pregnant? (16/16 positive responses to Family-of-Sentences variants)

(21) **–hechakuaa 'understand'**
Context: Ramona is visiting her mother in the hospital. She is sitting with her mother and overhears the doctor talking in the hallway to somebody. The doctor says:

O-hecha-kuaa o-sẽ-ma=ha Raul.
A3-see-know A3-leave-PRF=NMLZ Raul

'S/he/They understand that Raul has already left.'
Response question: Given what the doctor says, is he sure that Raul left already? (10/16 positive responses to Family-of-Sentences variants)

(22) **–juhu_1 'discover'**
Context: I need a small fan for my office. Luli says that she'll go to Salema later and so I ask her to look whether they sell small fans there. I don't see Luli again that day. Her mother says to me:

Lúli o-juhu o-ñe-vende=ha ventilador michĩ-va Saléma-pe.
Luli A3-discover A3-JE-sell=NMLZ fan small-REL Salema-PE

'Luli discovered that small fans are sold at Salema.'
Response question: Given what Luli's mother says, are small fans sold at Salema? (0/16 positive responses to Family-of-Sentences variants)

(23) a. **–mombe'u_1 'confess, tell'**
Context: Raul is my neighbor. He lives alone and doesn't have friends or family nearby. When he is sick, I bring him food. I am talking to his son who recently talked to his father on the phone. The son says to me:

Raúl o-mombe'u chéve ha'e hasy=ha.
Raul A3-confess/tell pron.1sg.NONAG pron.3.AG B3sg.sick=NMLZ

'Raul told me that he is sick.'
Response question: Given what the son tells me, is he affirming that his father is sick? (9/16 positive responses to Family-of-Sentences variants)

b. **–mombe'u_2 'confess, tell'** Context: Yesterday I bought ice cream for my husband's birthday on Sunday. But when I checked on it again today, a large part was already missing. I'm going to send whoever ate the ice cream to the store to buy more ice cream. My daughter says about my son, Manuel:

Manuel o-mombe'u ore-rú-pe ho'u=ha-gue eládo.
Manuel A3-confess/tell B1pl.excl-father-PE A3.eat=NMLZ-NOM.TERM ice.cream

'Manuel confessed to our father that he ate the ice cream.'
Response question: Given what my daughter says, will I send Manuel to buy more ice cream? (12/16 positive responses to Family-of-Sentences variants)

c. **–mombe'u_3 'confess, tell'**
Context: My daughter Laura has a boyfriend. I worry about her and her boyfriend but she won't tell me anything about what's going on anymore. I hear my other daughter talk on the phone. She says:

Láura o-mombe'u i-kichihá-me hy'e guasu=ha.
Laura A3-confess/tell B3-boyfriend-PE stomach big=NMLZ

'Laura confessed to her boyfriend that she is pregnant.'
Response question: Given what my daughter says about her sister Laura, is she sure that Laura is pregnant? (16/16 positive responses to Family-of-Sentences variants)

(24) **–'e 'say'**
Context: Raul is my neighbor. He lives alone and doesn't have friends or family nearby. When he is sick, I bring him food. I am talking to his son who recently talked to his father on the phone. The son says to me:

Raúl he'i ha'e hasy=ha.
Raul A3.say pron.3.AG B3sg.sick=NMLZ

'Raul said that he is sick.'
Response question: Given what the son tells me, is he affirming that his father is sick? (8/16 positive responses to Family-of-Sentences variants)

(25) **–ha'arõ 'hope'**
Context: Miriam lives with her sister. The two have a key to their home, and nobody else has one. Today Miriam lost her key and she is on her way home. Her friend, who knows that Miriam lost her key and that she is on her way home, says to me:

Míriam o-ha'arõ iñ-ermána o-ĩ=ha hóga-pe.
Miriam A3-hope B3-sister A3-be=NMLZ B3.house-PE

'Miriam hopes that her sister is at home.'
Response question: Given what Miriam's friend said, is she affirming that Miriam's sister is at home? (1/16 positive responses to Family-of-Sentences variants)

B Experiment stimuli

The experiment stimuli were recordings of the sentences in (26) to (28). The recordings can be found in the GitHub repository mentioned in footnote 9. Each example also identifies the content whose projectivity was explored by providing the clausal complement of the response question *Según lo que preguntó Magda, que certeza tiene de que...?* 'According to what Martha asked, is she certain that...?'.

(26) **Familiarization items**

1 Abél kichihá=pa o-jeroky-kuaa?
Abel girlfriend=Q A3-dance-know
'Can Abel's girlfriend dance?'
Response question: ...*Abel tiene una novia?* '...Abel has a girlfriend?'

2 Ramón=pa o-ñembyasy o-pe=ha hetyma?
Ramon=Q A3-regret A3-break=NMLZ B3.leg
'Does Ramon regret that he broke his leg?'
Response question: ...*Ramon rompió su pierna?* '...Ramon broke his leg?'

3 Abél=pa o-guereko kichiha?
Abel=Q A3-have girlfriend
'Does Abel have a girlfriend?'
Response question: ...*Abel tiene una novia?* '...Abel has a girlfriend.'

(27) **Main clause control items**

A1 Nína=pa o-gueru kuri peteĩ yvoty?
Nina=Q A3-bring past one flower
'Did Nina bring a flower?'
Response question: ...*Nina trajo una flor?* '...Nina brought a flower?'

A2 Chrístian=pa o-hayhu Katalína-pe?
Christian=Q A3-love Katalina-PE

‘Does Christian love Katalina?’
Response question: ...*Christian ama a Katalina?* ‘...Christian loves Katalina?’

A3 Fátima=pa o-kosiná-ma mbeju?
Fatima=Q A3-cook-PRF mbeju
‘Did Fatima already make mbeju?’
Response question: ...*Fatima cocinó mbeju?* ‘...Fatima already made mbeju?’

B1 Daniél=pa oi-kytĩ peteĩ yvyramáta?
Daniel=Q A3-cut one tree
‘Did Daniel cut a tree?’
Response question: ...*Daniel corto un arbol?* ‘...Daniel cut a tree?’

B2 Blánca=pa o-lee hína peteĩ líbro?
Blanca=Q A3-read PROG one book
‘Is Blanca reading a book?’
Response question: ...*Blanca esta leyendo un libro?* ‘...Blanca is reading a book?’

B3 Pédro=pa o-karú-ma?
Pedro=Q A3-eat-PRF
‘Has Pedro eaten already?’
Response question: ...*Pedro ya comió?* ‘...Pedro has eaten already?’

(28) **Target items**

a. Non-restrictive relative clauses

1 Cármen=pa o-visitá-ta i-sý-pe Brasíl-pe oi-kó-va?
Carmen=Q A3-visit-PROSP B3-mother-PE Brazil-PE A3-live-REL
‘Is Carmen going to visit her mother, who lives in Brazil?’
Response question: ...*la mama de Carmen vive en Brasil?* ‘...Carmen’s mother lives in Brazil?’

2 Sándro=pa o-henói-ta i-túva-pe campaña-pe o-mba’apó-va?
Sandro=Q A3-call-PROSP A3-father-PE countryside-PE A3-work-REL
‘Is Sandro going to call his father, who works in the countryside?’
Response question: ...*el papa de Sandro trabaja en la campaña?* ‘...Sandro’s father works in the countryside?’

b. Possessive noun phrases

1 Sándra lóro=pa hasy?
Sandra parrot=Q B3.sick
‘Is Sandra’s parrot sick?’
Response question: ...*Sandra tiene un loro?* ‘...Sandra has a parrot?’

2 Mário ryguasú=pa o-kuru?
Mario chicken=Q A3-brood
‘Is Mario’s chicken brooding?’
Response question: ... *Mario tiene una gallina?* ‘...Mario has a chicken?’

c. Items with *=nte* ‘only’:

1 Láura=nte=pa o-gana?
Laura=only=Q A3-win
‘Did only Laura win?’
Response question: ...*Laura ganó?* ‘...Laura won?’

2 Márko=nte=pa ou?
Marko=only=Q A3.come
'Did only Marko come?'
Response question: *...Marko vino?* '...Marko came?'

d. Items with change-of-state expression:

1 Éva=pa nd-o'u-vé-i-ma caña?
Eva=Q NEG-A3.eat-more-NEG-PRF caña
'Has Eva stopped drinking caña?'
Response question: *...Eva ha tomado caña?* '...Eva has drunk caña?'

2 Felípe=pa nd-o-pita-vé-i-ma?
Felipe=Q NEG-A3-smoke-more-NEG-PRF
'Has Felipe stopped smoking?'
Response question: *...Felipe ha fumado?* '...Felipe has smoked?'

e. Items with *haimete* 'almost':

1 Haimeté=pa o-menda Júlia?
almost=Q A3-marry Julia
'Did Julia almost get married?'
Response question: *...Julia sigue soltera?* '...that Julia is still single?'

2 Haimeté=pa o-mano Júlio?
almost=Q A3-die Julio
'Did Julio almost die?'
Response question: *...Julio sigue vivo?* '...Julio is still alive?'

f. Items with *–vy'a* 'happy':

1 Dáni=pa o-vy'a o-ñemboki=ha iñ-ermána?
Dani=Q A3-happy A3-in.love=NMLZ A3-sister
'Is Dani happy that her sister is in love?'
Response question: *...la hermana de Dani esta enamorada?* '...Dani's sister is in love?'

2 Diégo=pa o-vy'a na-i-ñembyahýi-ri=ha iñ-ermáno?
Diego=Q A3-happy NEG-B3-hungry-NEG=NMLZ B3-brother
'Is Diego happy that his brother isn't hungry?' (for some speakers also: '...isn't stingy?')
Response question: *...el hermáno de Diego no es mezquino?* '...that Diego's brother isn't stingy?'

g. Items with *–hecha* 'see':

1 Liliána=pa o-hecha i-vesína o-sẽ=ha tupa'ó-gui?
Liliana=Q A3-see B3-neighbor A3-go.out=NMLZ church-ABL
'Did Liliana see her neighbor go out of the church?'
Response question: *...la vecina de Liliana salió de la iglesia?* '...Liliana's neighbor go out of the church?'

2 Raúl=pa o-hecha i-vesíno oi-ke=ha peteĩ otél-pe?
Raul=Q A3-see B3-neighbor A3-enter=NMLZ one hotel-PE
'Did Raul see his neighbor enter a hotel?'
Response question: *...el vecino de Raul entró a un hotel?* '...Raul's neighbor entered a hotel?'

h. Items with *–hecha-uka* (see-CAUS) 'show, reveal':

1 Beníta=pa o-hecha-uka i-rika=ha?
Benita=Q A3-see-CAUS B3-rich=NMLZ
'Did Benita show/reveal that she is rich?'
Response question: *...Benita es rica?* '...Benita is rich?'

2 Róque=pa o-hecha-uka i-mboriahu=ha?
Roque=Q A3-see-CAUS B3-poor=NMLZ
'Did Roque show/reveal that he is poor?'
Response question: *...Roque es pobre?* '...Roque is poor?'

i. Items with *–hecha-kuaa* (see-know) 'understand, realize':

1 Juána=pa o-hecha-kuaa i-túva i-pochy=ha?
Juana=Q A3-see-know B3-father B3-angry=NMLZ
'Does Juana understand/realize that her father is angry?'
Response question: *...el papa de Juana esta enojado?* '...Juana's father is angry?'

2 Gustávo=pa o-hecha-kuaa i-sy iñ-angekoi=ha?
Gustavo=Q A3-see-know B3-mother B3-annoyed=NMLZ
'Does Gustavo understand/realize that his mother is annoyed?'
Response question: *...la mama de Gustavo esta molesta?* '...that Gustavo's mother is annoyed?'

j. Items with *–hendu* 'hear':

1 Tína=pa o-hendu ij-ekípo o-perde=ha-gue?
Tina=Q A3-hear B3-team A3-lose=NMLZ-NOM.TERM
'Did Tina hear that her team lost?'
Response question: *...el ekipo de Tina perdió?* '...Tina's team lost?'

2 Enríque=pa o-hendu i-kolektívo itinerário o-je-guerova=ha-gue?
Enrique=Q A3-hear B3-bus itinerary A3-JE-change=NMLZ-NOM.TERM
'Did Enrique hear that the itinerary of his bus changed?'
Response question: *...el itinerario del colectivo se cambió?* '...the itinerary of the bus changed?'

k. Items with *–juhu* 'discover':

1 Lídia=pa o-juhu i-juky=ha i-mburuvicha?
Lidia=Q A3-discover B3-salt=NMLZ B3-boss
'Did Lidia discover that her boss is nice?'
Response question: *...el jefe de Lidia es simpatico?* '...that Lidia's boss is nice?'

2 Páblo=pa o-juhu horý=ha i-mbo'ehára?
Pablo=Q A3-discover B3.nice=NMLZ B3-teacher
'Did Pablo discover that his teacher is nice?'
Response question: *...la maestra de Pablo es amable?* '...Pablo's teacher is nice?'

l. Items with *–topa* 'discover':

1 Clára=pa o-topa i-ména o-kopi=ha-gue kapi'ipe?
Clara=Q A3-discover B3-husband A3-cut=NMLZ-NOM.TERM meadow
'Did Clara discover that her husband cut the meadow?'
Response question: *... el esposo de Clara corpió el pasto?* '...Clara's husband cut the meadow?'

2 Cárlo=pa o-topa iñ-angiru o-joka=ha-gue peteĩ térmo?
Carlo=Q A3-discover B3-friend A3-break=NMLZ-NOM.TERM one thermos
'Did Carlos discover that his friend broke a thermos?'
Response question: *...el amigo de Carlo rompió un termo?* '...Carlos' friend broke a thermos?'

m. Items with *–kuaa* 'know':

1 Elvíra=pa oi-kuaa i-memby o-jahu=ha-gue río-pe?
Elvira=Q A3-know B3-son A3-bathe=NMLZ-NOM.TERM river-PE
'Does Elvira know that her son bathed in the river?'
Response question: *...el hijo de Elvira se baño en el rio?* '...Elvira's son bathed in the river?'

2 Darío=pa oi-kuaa i-tio o-ho=ha-gue videojuégo-pe?
Dario=Q A3-know B3-uncle A3-go=NMLZ-NOM.TERM video.game.store-PE
'Does Dario know that his uncle went to the video game store?'
Response question: *...el tio de Darío se fue al videojuego?* '...Dario's uncle went to the video game store?'

n. Items with *–kuaa-uka* (know-CAUS) 'point out':

1 Mírna=pa oi-kuaa-uka oi-ko=ha i-sý-ndi?
Mirna=Q A3-know-CAUS A3-live=NMLZ B3-mother-with
'Did Mirna point out that she lives with her mother?'
Response question: *...Mirna vive con su mama?* '...Mirna lives with her mother?'

2 Síxto=pa oi-kuaa-uka o-guereko=ha peteĩ kabaju?
Sixto=Q A3-know-CAUS A3-have=NMLZ one horse
'Did Sixto point out that he has a horse?'
Response question: *...Sixto tiene un caballo?* '...Sixto has a horse?'

o. Items with *–mombe'u* 'confess, tell':

1 Rósa=pa o-mombe'u hy'e guasu=ha-gue?
Rosa=Q A3-confess/tell B3.stomach big=NMLZ-NOM.TERM
'Did Rosa confess that she was pregnant?'
Response question: *...Rosa estaba embarazada?* '...that Rosa was pregnant?'

2 Nélson=pa o-mombe'u o-guenohẽ=ha-gue úno matemátika-pe?
Nelson=Q A3-confess/tell A3-get=NMLZ-NOM.TERM one math-PE
'Did Nelson confess that he got a one in math?'
Response question: *...Nelson sacó un uno en matematika?* '...Nelson got a one in math?'

p. Items with *–'e* 'say':

1 Míriam=pa he'i i-kuriosa=ha i-vesína?
Miriam=Q A3.say B3-nosy=NMLZ B3-neighbor
'Did Miriam say that her neighbor is nosy?'
Response question: *...la vesina de Miriam es curiosa?* '...Miriam's neighbor is nosy?'

2 Ricárdo=pa he'i i-kaigue=ha i-vesíno?
Ricardo=Q A3.say B3-exhausted=NMLZ B3-neighbor
'Did Ricardo say that his neighbor is exhausted?'
Response question: *...el vesino de Ricardo esta fatigado?* '...Ricardo's neighbor is exhausted?'

q. Items with *–mo'ã* 'think':

1 Dóra=pa oi-mo'ã Juã o-ñe-monde=ha hína?
Dora=Q A3-think Juan A3-JE-dress=NMLZ PROG
'Does Dora think that Juan is getting dressed?'
Response question: *...Juan se esta vistiendo?* '...Juan is getting dressed?'

2 Davíd=pa oi-mo'ã iñ-ermáno o-jahu=ha hína?
David=Q A3-think B3-brother A3-bathe=NMLZ PROG
'Does David think that his brother is bathing/taking a shower?'
Response question: *...el hermano de David se esta bañado?* '...that David's brother is bathing/taking a shower?'

C Experiment blocks A and B

Table A1 shows the order of the items in blocks A and B in the experiment reported on in section 2.3. The items are identified by the labels in (26) to (28) above.

item #	Block A	item #	Block B
1	control A1	21	control B1
2	*–vy'a* 'happy' 1	22	*–vy'a* 'happy' 2
3	*–mo'ã* 'think' 2	23	*–mo'ã* 'think' 1
4	possessive NP 1	24	possessive NP 2
5	*–kuaa-uka* 'point out' 2	25	*–kuaa-uka* 'point out' 1
6	*haimete* 'almost' 1	26	*haimete* 'almost' 2
7	*–hecha* 'see' 2	27	*–hecha* 'see' 1
8	*–mombe'u* 'confess, tell' 1	28	*–mombe'u* 'confess, tell' 2
9	*–topa* 'find out' 2	29	*–topa* 'find out' 1
10	control A2	30	control B2
11	*–hecha-uka* 'show, reveal' 1	31	*–hecha-uka* 'show, reveal' 2
12	*–kuaa* 'know' 2	32	*–kuaa* 'know' 1
13	change of state 1	33	change of state 2
14	*–hendu* 'hear' 2	34	*–hendu* 'hear' 1
15	*=nte* 'only' 1	35	*=nte* 'only' 2
16	*–juhu* 'discover' 2	36	*–juhu* 'discover' 1
17	*–hecha-kuaa* 'understand, realize' 1	37	*–hecha-kuaa* 'understand, realize' 2
18	*he'i* 'say' 2	38	*he'i* 'say' 1
19	NRRC 1	39	NRRC 2
20	control A3	40	control B3

Table A1: Items in blocks A and B

References

Abrusán, Márta. 2011. Predicting the presuppositions of soft triggers. *Linguistics & Philosophy* 34:491–535.

Abrusán, Márta. 2016. Presupposition cancellation: Explaining the 'soft-hard' trigger distinction. *Natural Language Semantics* 24:165–202.

Abusch, Dorit. 2002. Lexical alternatives as a source of pragmatic presupposition. In *Semantics and Linguistic Theory (SALT) XII*, pages 1–19. Ithaca, NY: CLC Publications.

Abusch, Dorit. 2010. Presupposition triggering from alternatives. *Journal of Semantics* 27:37–80.

Anand, Pranav and Valentine Hacquard. 2014. Factivity, belief and discourse. In *The Art and Craft of Semantics: A Festschrift for Irene Heim*, pages 69–90. MIT Working Papers in Linguistics.

AnderBois, Scott, Adrian Brasoveanu, and Robert Henderson. 2015. At-issue proposals and appositive impositions in discourse. *Journal of Semantics* 32:93–138.

AnderBois, Scott and Robert Henderson. 2015. Linguistically established discourse context: Two case studies from Mayan languages. In R. Bochnak and L. Matthewson, eds., *Methodologies in Semantic Fieldwork*, pages 207–232. Oxford: Oxford University Press.

Beaver, David and Brady Clark. 2008. *Sense and Sensitivity: How Focus Determines Meaning*. Oxford: Wiley-Blackwell.

Beaver, David and Emiel Krahmer. 2001. A partial account of presupposition projection. *Journal of Logic, Language, and Information* 10:147–182.

Bochnak, M. Ryan and Lisa Matthewson, eds. 2015. *Methodologies in Semantic Fieldwork*. Oxford: Oxford University Press.

Boersma, Paul and David Weenink. 2016. Praat: Doing phonetics by computer [computer program]. Version 6.0.19, retrieved 13 June 2016 from http://www.praat.org/.

Chemla, Emmanuel. 2009. An experimental approach to adverbial modification. In U. Sauerland and K. Yatsushiro, eds., *Semantics and Pragmatics: From Experiment to Theory*. New York: Palgrave Macmillian.

Chierchia, Gennaro and Sally McConnell-Ginet. 1990. *Meaning and Grammar*. Cambridge, MA: MIT Press.

Christensen, Rune Haubo Bojesen. 2013. "ordinal": Regression models for ordinal data. R package. Version ordinal_2015.6-28.

Cohen, Jacob. 1988. *Statistical power analysis for the behavioral sciences*. New Jersey: Lawrence Erlbaum, 2nd edn.

Estigarribia, Bruno. 2017. A grammar sketch of Paraguayan Guaraní. In B. Estigarribia and J. Pinta, eds., *Guaraní Linguistics in the 21st Century*, pages 7–85. Leiden: Brill Publishing.

Fasoli-Wörmann, Daniela. 2002. *Sprachkontakt und Sprachkonflikt in Paraguay*. Frankfurt: Peter Lang.

Gregores, Emma and Jorge A. Suárez. 1967. *A Description of Colloquial Guaraní*. The Hague: Mouton de Gruyter.

Grice, Paul. 1975. Logic and conversation. In P. Cole and J. Morgan, eds., *Syntax and Semantics 3: Speech Acts*, pages 64–75. New York: Academic Press.

Gutzmann, Daniel. 2015. *Use-Conditional Meaning: Studies in Multidimensional Semantics*. Oxford: Oxford University Press.

Heim, Irene. 1983. On the projection problem for presuppositions. In M. Barlow, D. Flickinger, and M. Westcoat, eds., *West Coast Conference on Formal Linguistics (WCCFL) 2*, pages 114–125.

Horn, Lawrence R. 1996. Exclusive company: Only and the dynamics of vertical inference. *Journal of Semantics* 13:1–40.

Horn, Laurence R. 2002. Assertoric inertia and NPI licencing. In *Proceedings of the 38th Annual Meeting of the Chicago Linguistic Society: Parasession on Negation and Polarity*, vol. 38, pages 55–82.

Hothorn, Torsten, Frank Bretz, and Peter Westfall. 2008. Simultaneous inference in general parametric models. *Biometrical Journal* 50(3):346–363.

Kadmon, Nirit. 2001. *Formal Pragmatics*. Oxford: Blackwell.

Karttunen, Lauri. 1971. Some observations on factivity. *Papers in Linguistics* 4:55–69.

Levinson, Stephen C. and E. Annamalai. 1992. Why presuppositions aren't conventional. In R. N. Srivastava, ed., *Language and Text: Studies in Honour of Ashok R. Kelkar*, pages 227–242. Delhi: Kalinga Publications.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70:369–415.

Murray, Sarah. 2014. Varieties of update. *Semantics & Pragmatics* 7(2):1–53.
Potts, Christopher. 2005. *The Logic of Conventional Implicatures*. Oxford: OUP.
Potts, Chris. 2007. The expressive dimension. *Theoretical Linguistics* 33:165–197.
R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
Roberts, Craige. 2011. Only: A case study in projective meaning. In B. Partee, M. Glanzberg, and J. Skilters, eds., *Formal Semantics and Pragmatics: Discourse, Context, and Models, The Baltic International Yearbook of Cognition, Logic and Communication*, vol. 6. New Prarie Press.
Romoli, Jacopo. 2015. The presuppositions of soft triggers are obligatory scalar implicatures. *Journal of Semantics* 32:173–291.
Schlenker, Philippe. 2010. Local contexts and local meanings. *Philosophical Studies* 151:115–142.
Simons, Mandy. 2001. On the conversational basis of some presuppositions. In *Semantics and Linguistics Theory (SALT) XI*, pages 431–448. Ithaca, NY: CLC Publications.
Simons, Mandy, David Beaver, Craige Roberts, and Judith Tonhauser. 2017. The Best Question: Explaining the projection behavior of factive verbs. *Discourse Processes* 3:187–206.
Smith, E. Allyn and Kathleen Currie Hall. 2011. Projection diversity: Experimental evidence. In *Proceedings of the 2011 ESSLLI Workshop on Projective Content*, pages 156–170.
Spector, Benjamin and Paul Egré. 2015. A uniform semantics for embedding interrogatives: *An* answer, not necessarily *the* answer. *Synthese* 192:1729–1784.
Stewart, Andrew. 2017. Jopara and the Spanish-Guarani language continuum in Paraguay: Considerations in linguistics, education, and literature. In B. Estigarribia and J. Pinta, eds., *Guaraní Linguistics in the 21st Century*, pages 379–416. Leiden: Brill Publishing.
Tonhauser, Judith, David Beaver, and Judith Degen. 2018. How projective is projective content? Gradience in projectivity and at-issueness. *Journal of Semantics* 35:495–542.
Tonhauser, Judith, David Beaver, Craige Roberts, and Mandy Simons. 2013. Toward a taxonomy of projective content. *Language* 89:66–109.
Tonhauser, Judith, Marie-Catherine de Marneffe, Shari R. Speer, and Jon Stevens. 2019. On the information structure sensitivity of projective content. In *Sinn und Bedeutung 23*.
van der Sandt, Rob. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9:333–377.
Velázquez-Castillo, Maura. 2004a. Guaraní (Tupí-Guaraní). In G. Booij, C. Lehmann, J. Mugdan, and S. Skopeteas, eds., *Morphology: An International Handbook on Inflection and Word-Formation*, vol. 2, pages 1421–1432. Berlin: Walter de Gruyter.
Velázquez-Castillo, Maura. 2004b. Serial verb constructions in Paraguayan Guaraní. *International Journal of American Linguistics* 70(2):187–213.
von Gleich, Utta. 1993. Paraguay – Musterland der Zweisprachigkeit? *Quo vadis Romania* 1:19–30.
Xue, Jingyang and Edgar Onea. 2011. Correlation between projective meaning and at-issueness: An empirical study. In *Proceedings of the 2011 ESSLLI Workshop on Projective Content*, pages 171–184.