

YOUNG ADULTS' INFERENTIAL READING SKILLS – RECOGNISING IMPLICIT AND UNSTATED CONTENT IN ITALIAN L1 AND ENGLISH L2

*Sara Gesuato*¹, *Elena Pagliarini*², *Emanuela Sanfelici*³

1. INTRODUCTION

Understanding a written text requires processing it at three levels. First, it involves decrypting its code, that is, recognising its graphic symbols and associating these with their conventional meanings (e.g. that the word *dog* stands for the concept 'dog'). Second, understanding a written text involves making sense of what is overtly expressed, its literal content. This includes propositional meaning (i.e. descriptions and representations of states of affairs). It may also include attitudinal meaning (i.e. the manifestation of the writer's state of mind like surprise encoded in *Oh*) and/or procedural meaning (i.e. the signalling of the writer's communicative intention relevant to the proposition conveyed like asserting in a declarative sentence). Grasping propositional, attitudinal and procedural meaning that is explicitly encoded is literal reading. Finally, a full understanding of a written text involves retrieving its implied meanings (i.e. what is indirectly expressed) by resorting to contextual clues, such as identifying relationships between meaning units, pinning down the writer's attitude (e.g. their level of confidence as revealed by their use of modal verbs) or intention (i.e. if they use an interrogative sentence to ask a question or request information), assigning reference, resolving ambiguity, and enriching explicit content with relevant details. This is inferential reading.

Although exchanging implied meanings is advantageous to the sender and the addressee, since more content is conveyed in less time, inferential reading is more demanding than literal reading (Alptekin and Erçetin, 2010: 207-209). First, it requires more cognitive effort, that is, engaging in more controlled processing; it also requires the production of novel knowledge from the text and the integration of the content from the text with one's previous knowledge. More importantly, the type of cognitive processing required is also not constant because of the varied nature of implied meanings: some are provided in addition to the literal meaning (e.g. presuppositions); others are conveyed in alternative to it (e.g. generalised implicatures); others are effectively meant to replace the literal meaning by contradicting it (e.g. irony).

Inferential reading may give rise to problems when implied meanings go unnoticed or are misinterpreted. This may happen when participants with different backgrounds have

¹ Università degli Studi di Padova.

² Universitat Autònoma de Barcelona.

³ Università degli Studi di Padova.

The authors are jointly responsible for the design of the study and the collection of data. The first author wrote Sections 1 and 3. The second author wrote Sections 2 and 6. The third author wrote Sections 4 and 5. The second author acknowledges financial support from the Spanish Ministerio de Ciencia e Innovación (Ramón y Cajal grant RYC2021-033969-I and project PID2022-138413NB-I00) and the Generalitat de Catalunya (grant 2021SGR00787).

different expectations about what should be explicitly communicated, when they find different aspects of the situation selectively salient to them, or when they have limited language proficiency (see Section 2).

In our study, we investigate how successfully L1 and L2⁴ speakers engage in inferential reading, when aiming to retrieve different kinds of meanings. In the rest of our paper, we provide a review of studies on inferential skills, focusing on L1 and L2 speakers (Section 2), outline our research method and specify our research questions (RQs; Section 3), and present (Section 4) and discuss (Section 5) our findings, drawing conclusions from them (Section 6).

2. LITERATURE REVIEW

Previous studies on the computation (i.e. processing) of pragmatic inferences in L2 speakers show contradictory evidence. Some studies suggest that L2 speakers can readily retrieve implied meanings (Feng, 2022; Feng and Cho, 2019; Lieberman, 2009; Miller *et al.*, 2016; Slabakova, 2010; Snape and Hosoi, 2018.), whereas others suggest that they may experience more difficulty than L1 speakers (Khorsheed *et al.*, 2022; Ko *et al.*, 2010; Mazzaggio *et al.*, 2021).

One theory put forward to explain the difficulties experienced by L2 learners in the derivation of implied meanings is the so-called Interface Hypothesis (Sorace, 2011; Sorace and Filiaci, 2006). This hypothesis posits that narrow syntactic properties should pose no acquisitional challenges to adult L2 speakers, but that difficulties arise for those phenomena at the interface between grammar and other cognitive systems, as well as between grammar and contextual variables (i.e. phenomena that involve contextual information external to syntax, i.e. discourse-pragmatic variables). In other words, linguistic phenomena involving the interaction between different cognitive domains are more challenging for L2 speakers than those encompassing internal interfaces, which involve the interaction between different aspects of the same cognitive domain. Evidence in favour of this hypothesis comes from studies showing over-extension of the use of overt subject pronoun in near-native L2 speakers of Italian (Sorace and Filiaci, 2006; Belletti *et al.*, 2007; see also Tsimpli *et al.*, 2004 and Valenzuela, 2006 for evidence in different languages and different linguistic constructions). Nevertheless, several studies suggest that neither deriving scalar implicatures⁵ (Feng and Cho, 2019; Lieberman, 2009; Miller *et al.*, 2016; Slabakova, 2010; Snape and Hosoi, 2018) nor computing (i.e. retrieving) presuppositions (Feng, 2022) is a problem for L2 speakers, thus challenging the claim of the Interface Hypothesis. Interestingly, a study by Anggraini and Sari (2023) showed that L1 English-speaking teachers of English at an Indonesian university used a variety of presuppositions to convey meaning and to facilitate communication with L2 learners, especially structural and counter-factual presuppositions.

Nevertheless, in some of the above studies, fine-grained differences between L1 and L2 speakers emerged. Feng (2022) found that L1 English speakers and L1-Mandarin Chinese L2-English speakers generated the inference of the presupposition trigger *stop* in affirmative and negated sentences at similar rates in both conditions, yet L2 speakers were

⁴ We use the term L2 to refer to both second and foreign languages.

⁵ The violation of the Maxim of Quantity – which states that speakers should be maximally informative – gives rise to a particular implicature, defined by Horn (1972; 1989) as scalar implicature. Horn introduces the notion of informativeness scale, in which sets of words are ranked by order of informativeness, from the weakest (least informative) to the strongest (most informative) (e.g. *or*, *and*, *some*, *many*, *most*, *all*, *might*, *must*).

significantly slower than L1 speakers in processing the inference. Furthermore, L2 speakers were less likely to suspend, that is, more likely to fail to retrieve, presuppositions than L1 speakers. Feng and Cho (2019) found no difference between L1 English speakers and L1-Mandarin Chinese L2-English speakers in the computation (i.e. derivation) of the direct scalar implicature associated with *sometimes* (\sim *not always*), but differences were found in the derivation of the indirect scalar implicature associated with *not always* (\sim *sometimes*), as L2 speakers computed it less frequently than L1 speakers (see also Taguchi, 2009 for the effect of differential degrees of implicitness in L2 speakers).

Research on L2 speakers' inferential skills has also investigated the processing of scalar implicatures. Many psycholinguistic studies showed that scalar implicature computation requires more processing time than logical (or literal) interpretation due to increased cognitive effort (Bott and Noveck, 2004; Breheny *et al.*, 2006; Dieussaert *et al.*, 2011; Huang and Snedeker, 2009; Noveck and Posada, 2003; Politzer-Ahles and Gwilliams, 2015). These findings support a contextual account of scalar implicatures whereby the pragmatic meaning is not automatically derived, as per the default approach (Levinson, 2000), but is instead arrived at after the logical meaning has been computed (Sperber and Wilson, 1987). Recently, the contextual view of scalar implicatures – according to which implicatures arise only if there is some contextual reason – has been supported by findings from studies carried out on L2 speakers. Mazzaggio *et al.* (2021) found that L1 speakers rejected, that is, dispreferred, more underinformative statements than L2 speakers when the sentences were orally presented under time constraints. These results are predicted under two conditions: first, when assuming that L2 speakers tested in the L2 language experience a greater cognitive load due to lower proficiency than L1 speakers tested in their L1; and second, when assuming that the pragmatic interpretations (i.e. those requiring integrating linguistic input with contextual details) are the non-default interpretations and thus require higher cognitive effort. It follows that proficiency might play a role in the computation of scalar inferences in the L2 population.

The above prediction has been born out in recent studies. Khorsheed *et al.* (2022) tested L1-Bahasa Malay L2-English speakers by means of a verification paradigm⁶, showing that participants with lower English proficiency were slower than participants with higher proficiency in computing the *some but not all* inference (i.e. scalar implicature) linked to *some* (see also Alptekin and Erçetin, 2010). When intermediate and advanced L2 participants were tested on the comprehension of scalar implicatures, no difference emerged between the two proficiency groups (Snape and Hosoi, 2018). Furthermore, working memory capacity also seems to play a significant role in inferential reading in L2 speakers, as it has been showed to positively correlate with inferential comprehension (Alptekin and Erçetin, 2010; Karimi and Naghdivand, 2017; Rai *et al.*, 2011). Finally, it has been shown that L2 vocabulary knowledge and basic decoding skills predict reading comprehension and lexical inferencing abilities in the L2 (Prior *et al.*, 2014).

However, while the literature focuses on the inferential skills of L2 learners – and those of populations with less-than-optimal cognitive capabilities (e.g. children: Bill *et al.*, 2016; Sbisà, 2007; elders: Domaneschi and Di Paola, 2019; Reinecke *et al.*, 2022; Atypical Development individuals: Bishop and Adams, 1992; Gough *et al.*, 2018) – it has so far neglected other populations like L1 speakers and Typical Development (TD) young adults; similarly scant are contrastive studies on L1 Italian and L2 English, studies examining inferential skills on extended stretches of discourse, and research investigating a range, rather than a single type, of inferential skills. We thus conducted a study to

⁶ A verification paradigm is a procedure for measuring reading comprehension. It usually involves having participants make judgments about a given sentence (e.g. *Some parrots are birds*), such as determining if it is true or false, and often recording their reaction times.

investigate how L1 and L2 TD adults perform in a reading comprehension task targeting various types of meanings.

3. METHOD

In an exploratory study, we compared and contrasted L1 and L2 TD speakers' ability to detect various types of meanings in a written narrative text. We limited our investigation to implied (i.e. presupposed and entailed) and unstated meanings. Implied meaning is meaning retrievable from explicit content via an inference based on (non-)linguistic evidence. More specifically, a presupposition is an underlying assumption, whose truth and validity are taken for granted, that ensures the logical acceptability of the proposition conveyed and thus the interpretability of the utterance. It is based on the meaning of words and structures and their conventional interpretation. For example, *My sister is at work* presupposes that the speaker/writer has a sister (see *My*), and *"I have been promoted to senior manager," beamed Tom*, presupposes that the speaker had a job before the promotion (see *promote*) and that he was happy about it (see *beam*). An entailment is instead the inescapable logical consequence of a proposition, which is enforced by lexical meanings. For example, *The coaster is under the glass* entails that the glass is on the coaster, and *"I have been promoted to senior manager," beamed Tom* entails that Tom was happy about it (see *beam*). Finally, unstated content is information that is neither overtly expressed nor recoverable through inferences. For example, from *Mary is my sister* it is possible to infer that Mary is female, but not how old she is (for details, see Sbisà, 2007).

The goal was to investigate whether the degree of accuracy with which information is recognised varies with: (a) the nativeness vs. non-nativeness of the readers, (b) the type of meanings to be retrieved (presuppositions vs entailments vs unstated content), and (c) the readers' level of L2 proficiency.

3.1. Design and material

We designed an Italian and an English version of an online questionnaire, administered through Google Forms⁷. We designed our instrument from scratch, rather than adapting one from the literature, since we wanted to study various kinds of inferential skills on an extended text. Each version included: an introductory statement; a few questions about the compiler's demographic data; a reading passage; 19 multiple-choice comprehension items focused on the retrieval of implied information (11 items on presuppositions, 4 on entailments) and the recognition of unstated content (4 items)⁸, and an optional open-ended question for possible comments.

The reading passage, titled "Angela", was a made-up story about a US⁸ citizen travelling to Germany for personal and professional reasons, and then moving back to the US (285 words in Italian and 292 words in English). It exemplified the classic structure of narratives as per the model by Labov (1972: 354). Therefore, it contained an Orientation, a Complicating action, a Peak, an Evaluation, and a Resolution.

⁷ The two versions of the questionnaire, the dataset and the R-script for the statistical analyses can be found on the Open Science Framework website:

https://osf.io/bhua5/?view_only=1c4ec07364bb4dbc8aa4f09a9be5e4e9.

⁸ Given that we tested inferential skills on a whole narrative text, we had to adapt our items to (the sequencing of) its content. For this reason, we had a different number of items across types of meanings.

The reading passage was presented first in its entirety, and then in short excerpts, each accompanied by one or more comprehension items in the form of statements. These had to be judged in terms of their accuracy by choosing one of the following options: *True, False, Not Given, Not Know*⁹, and also perceived level of difficulty by choosing one of the following options: *Very Easy, Somewhat Easy, Somewhat Hard, Very Hard*¹⁰.

For example, the text excerpt *She succeeded in being promoted and receiving a pay increase in less than 6 months* was associated with the item *Angela worked hard to climb up the career ladder*, which conveyed a presupposition, and which was supposed to be judged as 'True'. Instead, the excerpt *Of course, she would need time to re-adjust to her old-new life, but she was also looking forward to it. There was a lot of catching up to do. She had a lot to share. She was ready for a new phase of her life* was associated with the item *Angela was optimistic about her future*, which conveyed an entailment and was to be recognised as 'True'. Finally, the excerpt *She had been excited about this new adventure: she had a new job waiting for her and she was looking forward to putting her knowledge of German into practice* was associated with the item *Angela was very fluent in German*, which conveyed unstated content, and was thus to be classified as Not Given.

The comprehension items envisaged 9 True answers (2 from entailments, 7 from presuppositions), 6 False answers (2 from entailments, 4 from presuppositions) and 4 Not Given answers.

3.2. Participants

To recruit participants, we enlisted the help of colleagues from our and other universities in Italy. We asked them to invite their students to compile the questionnaire in their free time on a voluntary basis. They were randomly assigned to the English vs the Italian version if the last digit of their student ID number was an even vs odd number, respectively. In total, 108 students completed the questionnaire (54 per version). We excluded from further analysis the questionnaires of 10 people (5 per group) either because they did not state that Italian was (one of) their L1(s), or because they stated that, or failed to state whether, they had been diagnosed with some form of language impairment or learning disorder.

The participants' mean age was 21 years, 9 months (SD: 5 years, 9 months). They were enrolled in various degree courses, mainly in the humanities (Linguistics, Foreign Languages, Literature, Philosophy, History, Classics), and partly in the social sciences (i.e. Economics: 10; Political Science: 14; Peace and Conflict Studies: 1; Law: 1), or others (Biology: 1; Undeclared: 6). The compilers of the English questionnaire were asked to self-assess their English-L2 proficiency level following the Common European Framework of Reference for Languages: C2 native-like, C1 advanced, B2 high intermediate and B1 low intermediate¹¹. Table 1 reports the number of participants divided according to the level of proficiency they declared.

⁹ *Not Given* was to be used if the reader realised that a given piece of information was missing from the text, while *Not Know* was to be used if the reader was unsure as to which answer to choose between *True, False* and *Not Given*.

¹⁰ For reasons of space, this part of the data will not be considered.

¹¹ We did not test the participants' L2 English proficiency, since the questionnaire was already fairly long, a placement test would have taken additional time, and we could not reward students – financially or otherwise – for their participation.

Table 1. *English questionnaire participants' self-declared L2 proficiency levels*

Level of English-L2 proficiency	Number of participants (N=49)
Native-like	1
Advanced	15
High-intermediate	27
Low-intermediate	6

3.3. RQs and hypotheses

We formulated three RQs and hypotheses about the expected findings on the basis of the literature (Section 2).

RQ1: How accurate are L1 Italian speakers vs L1-Italian L2 English speakers at detecting implied (presupposed and entailed) meanings and recognising unstated meanings in a written narrative text?

Hypothesis 1a: L2 speakers will be less accurate than L1 speakers at deriving pragmatic inferences because these involve the interaction between different cognitive domains, a condition that has been reported to be challenging to L2 speakers.

Hypothesis 1b: L2 speakers will be as accurate as L1 speakers at deriving pragmatic inferences because the former have been known to readily derive implied meanings.

RQ2 Do meanings of different types affect the degree of accuracy with which they are retrieved in a written text?

Hypothesis 2: Yes, different types of meanings will pose different challenges to readers, since they require different kinds of inferences.

RQ3: Do different levels of English-L2 proficiency make a difference in detecting implied meanings and recognising unstated meanings in a written narrative text?

Hypothesis 3: L2 speakers with higher English proficiency will be more accurate than those with lower proficiency in detecting implied meanings and recognising unstated meanings in line with what reported about the computation of scalar inferences and reading comprehension abilities.

4. RESULTS

A total of 1,862 responses were collected and analysed, 931 in each version of the questionnaire. We first calculated the accuracy in retrieving meanings depending on the language of the questionnaire (RQ1). Figure 1 shows that the accuracy was slightly higher in the English texts (71%) than in the Italian ones (66%).

Figure 1. *Accurate responses in Italian vs. English (raw numbers and percentage values)*

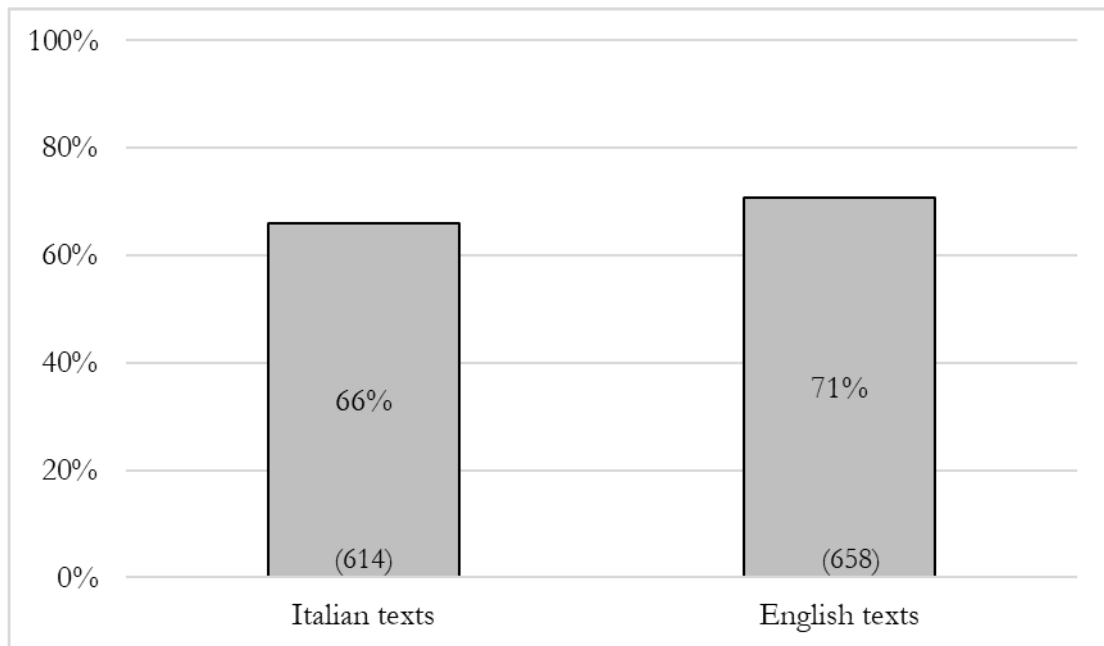
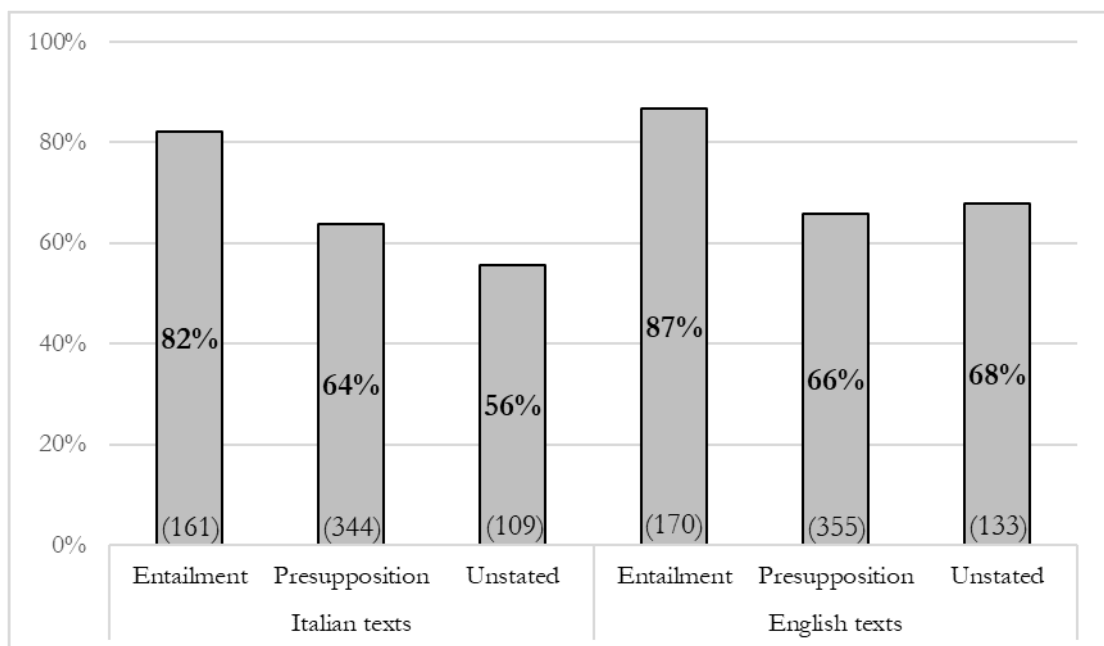


Figure 2. *Accurate responses in detecting entailed, presupposed, unstated meanings in Italian vs. English (raw numbers and percentage values)*



We then calculated participants' responses in detecting entailed, presupposed and unstated meanings in the Italian and the English questionnaires (RQ2). Figure 2 shows that the rate of accuracy was different across the three types of meanings considered. In the Italian questionnaires, the accuracy rate reached 82% in entailments, 64% in presuppositions and 56% in unstated meanings, respectively. In the English questionnaires, the accuracy rate was high in entailments, reaching 87%, while participants were similarly less accurate at detecting presuppositions and unstated meanings: 66% and

68% of the time, respectively. In addition, the accuracy rate was higher in the English than the Italian questionnaires for entailments, presuppositions and unstated meanings.

Participants' responses were fitted to a generalised mixed-effects logistic regression model in the statistical programming environment R (R Core Team, 2022). We used the *glmer* function (*lme4* package, Bates *et al.*, 2015) with the specification of the binomial family and the logit link function to run the analysis. We posited participants' responses (accurate vs. wrong) as our dependent variable, where value 1 was assigned to accurate responses and value 0 to wrong ones. As fixed factors, the model contained (a) the language of the text (2 levels: Italian vs. English), (b) the type of meanings to retrieve (3 levels: entailments vs. presupposition vs. unstated meanings) and (c) one interaction of language-by-type of meaning. Levels of the factors were all mean-centred, using orthogonal sum-to-zero contrasts. As for (a) language of the text, the contrast checked the difference between the Italian (coded +0,5) and the English (coded -0,5) version of the text. As for (b) type of meaning, the first contrast checked the difference between entailments (coded as +2/3) vs. presuppositions and unstated meanings (both coded as -1/3); instead, the second contrast checked the difference between presuppositions (coded as +0,5) and unstated meanings (coded -0,5), while entailments received code 0. Items and participants were set as random effects grouping intercepts in the model, including the factors Meanings as random slopes by participants. The model included the maximal structure that allowed the models to converge (Barr *et al.*, 2013). The model detected a significant effect of the factor Language and of the first contrast.

Table 2. *Fixed-effects estimates of the generalised linear mixed model*

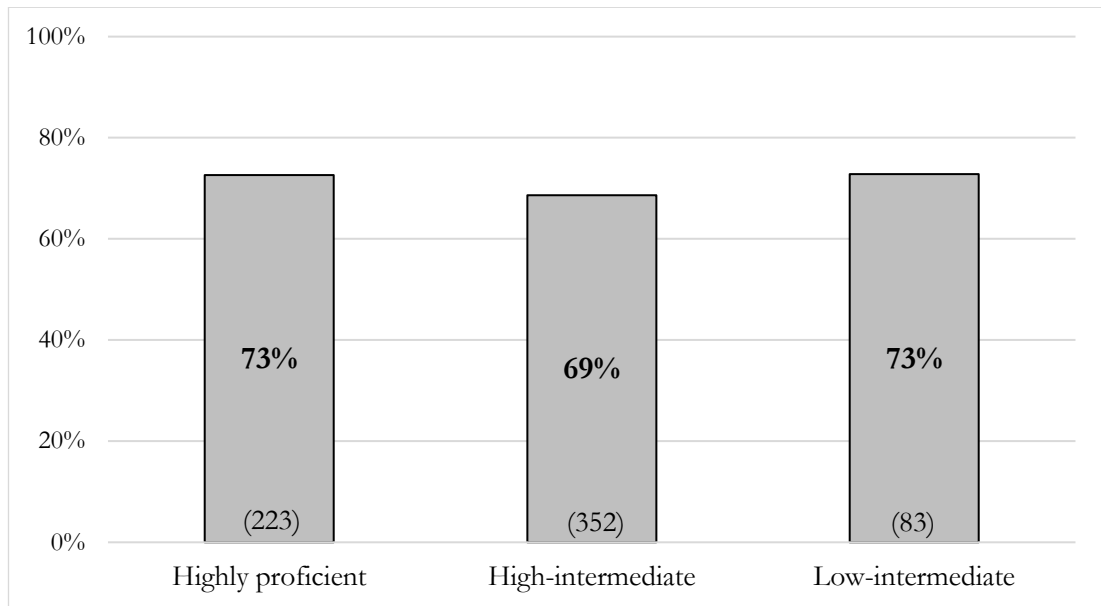
Fixed Factors	Estimate	SE	z	p
Intercept	1.14452	0.2724	4.202	2.65e-05***
Language	-0.37767	0.1328	-2.844	0.00446**
Contrast 1 (Entailment vs. Presupposition/Unstated)	1.34793	0.6080	2.217	0.02663 *
Contrast 2 (Presupposition vs. Unstated)	0.19989	0.5982	0.334	0.73827
Language * Contrast 1	-0.02518	0.3219	-0.078	0.93766
Language * Contrast 2	0.51198	0.2709	1.890	0.05879

(Full model summary: AIC= 2020.0; BIC= 2091.9; LogLik= -997.0; Dev= 1994.0)

The accuracy rate differed significantly between the two questionnaires. Participants gave significantly more accurate responses in English than in Italian. Contrast 1 indicates that participants were significantly more accurate with entailments than with the other types of meanings.

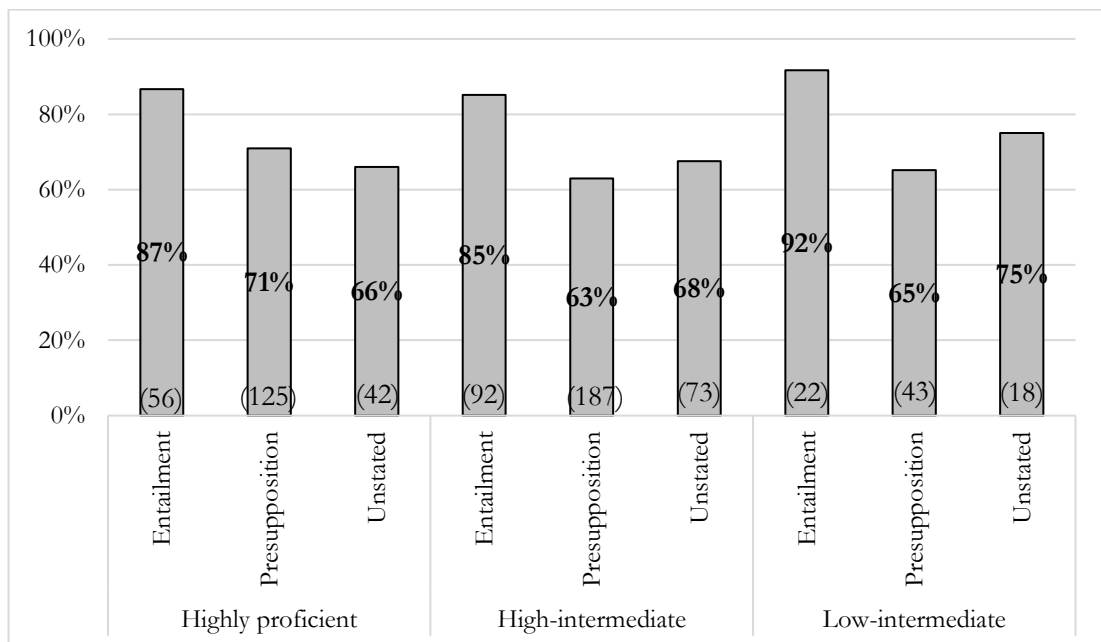
To address RQ3, we zoomed in on the responses collected with the English questionnaires. We calculated the percentages of participants' accurate responses arranged across the levels of self-declared English proficiency (Figure 3). Since there was only one participant in the native-like level, we grouped together the data about this speaker and the advanced-level speakers, in both the descriptive and the statistical analyses, referring to them collectively as highly proficient speakers.

Figure 3. Overall degree of accuracy in retrieving meanings across participants' levels of L2 proficiency (raw numbers and percentage values)



Highly proficient L2 speakers showed the same accuracy rate as low-intermediate ones, (73%), whereas high-intermediate speakers exhibited a lower accuracy rate (69%).¹²

Figure 4. Degree of accuracy in retrieving presupposed, entailed and unstated meanings across participants' levels of English proficiency (raw numbers and percentage values)



¹² For completeness' sake, we note that the accuracy rate of the only native-like L2 learner's responses (84%) was higher than the rates of the other participants.

Figure 4 shows the degree of accuracy reached by participants with different levels of L2 proficiency in detecting the three types of meanings. Whereas the highly proficient participants gave more accurate responses than the high-intermediate ones with entailments (87% vs. 85%) and presuppositions (71% vs. 63%), they were less accurate than the high-intermediate participants in unstated meanings (63% vs. 68%). In addition, the percentages of accurate responses by the low-intermediate L2 speakers were higher than those exhibited by the high-intermediate participants in the three types of meanings considered, i.e. entailments (92% vs. 85%), presuppositions (65% vs. 63%), and unstated meanings (75% vs. 68%). Likewise, the percentages of accurate responses by low intermediate L2 learners were higher than those exhibited by the highly proficient speakers in entailments and unstated meanings. Hence, the accuracy rates did not appear to decrease as a function of the participants' proficiency levels in all the three types of meanings¹³.

Participants' responses from the English questionnaires were fitted to a generalised mixed-effects logistic regression model with binomial family and logit link function. Participants' responses were set as our dependent variable, where value 1 was assigned to accurate responses and value 0 to wrong ones. As fixed factors, we posited (a) English-L2 proficiency levels (3 levels: highly proficient vs. high-intermediate vs. low-intermediate), (b) the type of pragmatic content (3 levels: entailments vs. presuppositions vs. unstated meanings) and (c) one interaction of level of proficiency-by-type of pragmatic meanings. Levels of the factors were all mean-centred, using orthogonal sum-to-zero contrasts. As for (a) L2 proficiency levels, the first contrast checked the difference between low-intermediate (coded +2/3) and the other two levels, high-intermediate and highly proficient (both coded -1/3). The second contrast checked the difference between highly proficient (coded +2/3) and the other two levels (both coded -1/3). As for (b) type of pragmatic content, the first contrast checked the difference between entailments (coded as +2/3) vs. presuppositions and unstated meanings (both coded as -1/3). The second contrast checked the difference between presuppositions (coded as +0,5) and unstated meanings (coded -0,5), while entailments received code 0. Items and participants were set as random effects grouping intercepts in the model, including the factor Meanings as random slopes by participants. The model included the maximal structure that allowed the models to converge (Barr *et al.*, 2013). The model detected no significant effect of the factors and of the interaction.

Table 3. *Fixed-effect estimates of the generalised linear mixed model*

Fixed Factors	Estimate	SE	z	p
Intercept	1.578544	0.436157	3.619	0.000296
Proficiency-Contrast 1 (Low-intermediate vs. High-intermediate/Highly proficient)	-0.429660	0.366085	-1.174	0.240531
Proficiency-Contrast 2 (Highly proficient vs. Low/High-intermediate)	0.191803	0.221662	0.865	0.386878
Meaning-Contrast 1 (Entailment vs. Presupposition/Unstated)	1.025682	0.982794	1.044	0.296652

¹³ For completeness' sake, we note that the accuracy rates of the only native-like L2 learner's responses were the following: 100% for entailments, 73% for presuppositions and 100% for unstated meanings.

Fixed Factors	Estimate	SE	z	p
Meaning-Contrast 2 (Presupposition vs. Unstated)	-0.043199	0.976359	-0.044	0.964709
Proficiency-Contrast 1 * Meaning-Contrast 1	-0.383327	0.853585	-0.449	0.653375
Proficiency-Contrast 2 * Meaning-Contrast 1	-0.005534	0.518402	-0.011	0.991482
Proficiency-Contrast 1 * Meaning-Contrast 2	0.329318	0.666553	0.494	0.621263
Proficiency-Contrast 2 * Meaning-Contrast 2	0.629099	0.449494	1.400	0.161642

(Full model summary: AIC= 933.2; BIC= 1010.6; LogLik= -450.6; Dev= 901.2)

5. DISCUSSION

RQ1 addressed how accurate L1-Italian speakers and L1-Italian L2-English speakers are at detecting implied meanings and recognising unstated meanings in a written narrative text. Two contrasting predictions were formulated based on previous findings (Section 2). On the basis of the results reported in Sorace and Filiaci (2006), Belletti *et al.* (2007) (see also Tsimpli *et al.*, 2004 and Valenzuela, 2006), L2 speakers were expected to be less accurate than L1 speakers at deriving pragmatic inferences because these involve the interaction between different cognitive domains, a condition that has been reported to be challenging to L2 speakers. Conversely, on the basis of other studies (Lieberman, 2009; Miller *et al.*, 2016; Slabakova, 2010; Snape and Hosoi, 2018), L2 speakers were predicted to be as accurate as L1 speakers at deriving pragmatic inferences. Our results did not provide support for either hypothesis. Overall, the accuracy in retrieving pragmatic meaning differed depending on the speakers' nativeness vs. non-nativeness in line with the first hypothesis. However, unexpectedly under both the first and the second hypothesis, the L2 speakers were significantly more accurate than the L1 speakers in retrieving implied meanings and recognising unstated meanings. One reason for this may be that students have more experience in reading comprehension tasks in their L2 than in their L1. Alternatively, they may have approached the task more carefully in the L2, expecting it would require more concentration, and this extra effort paid off.

RQ2 explored whether different types of meanings would differently affect the accuracy with which they would be recognised. As previous studies suggested that different types of pragmatic meanings may pose different challenges (Bott and Noveck, 2004; Breheny *et al.*, 2006; Dieussaert *et al.*, 2011; Huang and Snedeker, 2009; Noveck and Posada, 2003; Politzer-Ahles and Gwilliams, 2015), we expected a difference in the accuracy of participants' responses. Statistical analyses, indeed, revealed that the participants were significantly more accurate in entailments than in the other two types of meanings. Unstated meanings may be hard to identify because determining that something is not case may require a thorough processing of content details; presupposed meanings may be hard to identify because embedded in text segment that are not prominent in their formulation and/or position in the utterance, and thus are not salient to the reader; and entailments may be less hard to detect because they involve recognising obvious reformulations of content. As our design included a number of items unbalanced for the three types of meanings, future work will have to verify the solidity of this finding with a design controlled for this aspect.

RQ3 addressed whether different levels of English-L2 proficiency make a difference in detecting implied meanings and recognising unstated meanings in a written narrative text. According to recent studies on the computation of scalar inferences and reading comprehension skills (e.g. Mazzaggio *et al.*, 2021, Khorsheed *et al.*, 2022), we expected participants with higher L2 proficiency to be more accurate than those with lower proficiency in recognising implied and unstated meanings. In fact, the degree of accuracy in retrieving pragmatic meanings did not differ statistically across the levels of L2 proficiency. The discrepancy between our and previous findings may be due to the fact that the levels of L2 proficiency were unbalanced across participants: most declared they had a high-intermediate level of English proficiency, very few a low-intermediate level, and only one participant a native-like level. Alternatively, it may be that the participants' declared level of proficiency was not accurately assessed. Finally, it is possible that at a non-beginner level of proficiency, the ability to retrieve meanings in text is tied to cognitive factors that cut across language varieties (cf. Cummins's [1991], Model of Language Interdependence).

6. CONCLUSION

This work investigated the understanding of a written narrative text by L1 speakers of Italian and L1-Italian L2-English speakers. Our results showed that (a) L2 speakers were more accurate than L1 speakers in their overall reading comprehension, (b) that entailments were easier to retrieve than presuppositions or unstated information, and (c) that L2 speakers' performance did not seem to correlate with their level of proficiency. Result (b) is in line with previous findings that demonstrated some variation in computing different types of meaning (Feng, 2022; Feng and Cho, 2019). Conversely, results (a) and (c) are surprising, for which some speculative explanations were provided in Section 5.

This exploratory study suffers from some limitations, and thus provides only partial, non-conclusive, answers to the RQs addressed. Further research is needed to properly assess the solidity of our results. In future investigations, it would be important to control for a more homogeneous lexico-grammatical formulation of items (e.g. making sure that all presuppositions appear in thematic position in main clauses); to involve a more balanced number of participants across levels of L2 proficiency; to accurately determine participants' L2 proficiency with a placement test; to administer the test in a classroom, so as to ensure all participants do not have access to lexicographic resources; to measure the time taken to complete the questionnaire, as an additional factor in determining participants' performance; to include a balanced number of items across types of meanings; and to explore participants' conscious thought processes through post-test interviews.

REFERENCES

- Alptekin C., Erçetin G. (2010), “The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading”, in *Journal of Research in Reading*, 33, 2, pp. 206-219.
- Anggraini M. P., Sari R. N. (2023), “Teacher talk: a pragmatic analysis of presupposition in English teaching classroom”, in *20th AsiaTEFL-68th TEFLIN-5th iNELLTAL Conference (ASLATEFL 2022)*, Atlantis Press, Dordrecht, pp. 410-417.
- Barr D. J., Levy R., Scheepers C., Tily H. J. (2013), “Random effects structure for confirmatory hypothesis testing: keep it maximal”, in *Journal of Memory and Language*, 68, 3, pp. 255-278.
- Bates D., Mächler M., Bolker B. M., Walker S. C. (2015), “Fitting linear mixed-effects models using lme4”, in *Journal of Statistical Software*, 67, 1, pp. 1-48.
- Belletti A., Bennati E., Sorace A. (2007), “Theoretical and developmental issues in the syntax of subjects: evidence from near-native Italian”, in *Natural Language and Linguistic Theory*, 25, pp. 657-689.
- Bill C., Romoli J., Schwarz F., Crain S. (2016), “Scalar implicatures versus presuppositions: the view from acquisition”, in *Topoi*, 35, pp. 57-71.
- Bishop D. V. M., Adams C. (1992), “Comprehension problems in children with Specific Language Impairment: literal and inferential meaning”, in *Journal of Speech, Language, and Hearing Research*, 35, 1, pp. 119-129.
- Bott L., Noveck I. A. (2004), “Some utterances are underinformative: the onset and time course of scalar inferences”, in *Journal of Memory and Language*, 51, 3, pp. 437-457.
- Breheny R., Katsos N., Williams J. (2006), “Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences”, in *Cognition*, 100, 3, pp. 434-463.
- Cummins J. (1991), “Interdependence of first- and second-language proficiency in bilingual children”, in Bialystok E. (ed.), *Language processing in bilingual children*. Cambridge University Press, Cambridge, pp. 70-89.
- Dieussaert K., Verkerk S., Gillard E., Schaeken W. (2011), “Some effort for some: further evidence that scalar implicatures are effortful”, in *Quarterly Journal of Experimental Psychology*, 64, 12, pp. 2352-2367.
- Domaneschi F., Di Paola S. (2019), “The aging factor in presupposition processing”, in *Journal of Pragmatics*, 140, pp. 70-87.
- Feng S. (2022), “The computation and suspension of presuppositions by L1-Mandarin Chinese L2-English speakers”, in *Second Language Research*, 38, 4, pp. 737-763.
- Feng, S., Cho J. (2019), “Asymmetries between direct and indirect scalar implicatures in second language acquisition”, in *Frontiers in Psychology*, 10, 877: <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00877/full>.
- Gough Kenyon S. M., Palikara O., Lucas R. M. (2018), “Explaining reading comprehension in children with Developmental Language Disorder: the importance of elaborative inferencing”, in *Journal of Speech, Language and Hearing Research*, 61, 10, pp. 2517-2531.
- Horn L. R. (1972), “On the semantic properties of logical operators in English”, PhD thesis, University of California, Los Angeles.
- Horn L. R. (1989), *A Natural History of Negation*, The University of Chicago Press, Chicago.
- Huang Y. T., Snedeker J. (2009), “Online interpretation of scalar quantifiers: insight into the semantics-pragmatics interface”, in *Cognitive Psychology*, 58, 3, pp. 376-415.

- Karimi M. N., Naghdivand R. (2017), “Literal and inferential listening comprehension: the role of L1 vs. L2 auditory working memory capacity”, in *Journal of Modern Research in English Language Studies*, 4, 4, pp. 67-84.
- Khorsheed A., Md. Rashid S., Nimehchisalem V., Geok Imm L., Price J., Ronderos C. R. (2022), “What second-language speakers can tell us about pragmatic processing”, in *Plos one*, 17, 2:
<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0263724#abstract0>.
- Ko H., Ionin T., Wexler K. (2010), “The role of presuppositionality in the second language acquisition of English articles”, in *Linguistic Inquiry*, 41, 2, pp. 213-254.
- Labov W. (1972), *Language in the inner City*, University of Pennsylvania Press, Philadelphia.
- Levinson S. C. (2000). *Presumptive meanings: the theory of generalized conversational implicature*. MIT press, Cambridge, MA
- Lieberman M. (2009), “Necessary interpretation at the syntax/pragmatics interface: L2 acquisition of scalar implicatures, paper presented at *Workshop on Mind Context Divide: Language Acquisition and Interfaces of Cognitive Linguistic Modules*, University of Iowa, 30 April-2 May, 2009.
- Mazzaggio G., Panizza D., Surian L. (2021), “On the interpretation of scalar implicatures in first and second language”, in *Journal of Pragmatics*, 171, pp. 62-75.
- Miller D., Giancaspro D., Iverson M., Rothman J., Slabakova R. (2016), “Not just algunos, but indeed unos L2ers can acquire scalar implicatures in L2 Spanish”, in *Language Acquisition Beyond Parameters*, 51, pp. 125-145.
- Noveck I. A., Posada A. (2003), “Characterizing the time course of an implicature: an evoked potentials study”, in *Brain and Language*, 85, 2, pp. 203-210.
- Open Science Framework: www.osf.io.
- Politzer-Ahles S., Gwilliams L. (2015), “Involvement of prefrontal cortex in scalar implicatures: evidence from magnetoencephalography”, in *Language, Cognition and Neuroscience*, 30, 7, pp. 853-866.
- Prior A., Goldina A., Shany M., Geva E., Katzir T. (2014), “Lexical inference in L2: predictive roles of vocabulary knowledge and reading skill beyond reading comprehension”, in *Reading and Writing*, 27, pp. 1467-1484.
- R Core Team (2022), *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna: <https://www.R-project.org>.
- Rai M. K., Loschky L. C., Harris R. J., Peck N. R., Cook L. G. (2011), “Effects of stress and working memory capacity on foreign language readers’ inferential processing during comprehension”, in *Language Learning*, 61, 1, pp. 187-218.
- Reinecke R., Di Paola S., Domaneschi F., Fossard M. (2022), “Presupposition processing declines with age”, in *Cognitive Processing*, 23, pp. 479-502.
- Sbisà M. (2007), *Detto non detto. Le forme della comunicazione implicita*, Laterza, Bari-Roma.
- Slabakova R. (2010), “Scalar implicatures in second language acquisition”, in *Lingua*, 120, 10, pp. 2444-2462.
- Snape N., Hosoi H. (2018), “Acquisition of scalar implicatures: evidence from adult Japanese L2 learners of English”, in *Linguistic Approaches to Bilingualism*, 8, 2, pp. 163-192.
- Sorace A. (2011), “Pinning down the concept of “interface” in bilingualism”, in *Linguistic Approaches to Bilingualism*, 1, 1, pp. 1-33.
- Sorace A., Filiaci F. (2006), “Anaphora resolution in near-native speakers of Italian”, in *Second Language Research*, 22, 3, pp. 339-368.
- Sperber D., Wilson D. (1987), “Précis of relevance: communication and cognition”, in *Behavioral and Brain Sciences*, 10, 4, pp. 697-710.

- Taguchi N. (2009), “Corpus-informed assessment of comprehension of conversational implicatures in L2 English”, in *TESOL Quarterly*, 43, 4, pp. 738-749.
- Tsimpli T., Sorace A., Heycock C., Filiaci F. (2004), “First language attrition and syntactic subjects: a study of Greek and Italian near-native speakers of English, in *International Journal of Bilingualism*, 8, pp. 257-277.
- Valenzuela E. (2006), “L2 end state grammars and incomplete acquisition of the Spanish CLLD constructions”, in Slabakova R., Montrul S., Prévost P. (eds.), *Inquiries in Linguistic Development: in Honor of Lydia White*, John Benjamins, Amsterdam, pp. 283-304.

