



**HAL**  
open science

# L2 French learners' comprehension of non-literal language: A replication study on the comprehension of three different types of indirect meaning

Laëtitia Knight

► **To cite this version:**

Laëtitia Knight. L2 French learners' comprehension of non-literal language: A replication study on the comprehension of three different types of indirect meaning. Linguistics. Université Côte d'Azur, 2023. English. NNT: 2023COAZ2010 . tel-04129194

**HAL Id: tel-04129194**

**<https://theses.hal.science/tel-04129194>**

Submitted on 15 Jun 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE DE DOCTORAT

*La compréhension du langage non-littéral chez  
les apprenants de FLE: une étude de répliation  
sur la compréhension de trois types d'énoncés  
au sens indirect.*

L2 French learners' comprehension of non-literal  
language: a replication study on the comprehension  
of three different types of indirect meaning

**Laetitia JARLOT KNIGHT**

UMR 7320: Base, Corpus, Langage

**Présentée en vue de l'obtention  
du grade de docteur en Sciences du Langage  
d'Université Côte d'Azur**  
**Dirigée par :** Amanda EDMONDS, Professeure  
des Universités, Université Côte d'Azur  
**Soutenue le :** 17 mars 2023

**Devant le jury, composé de :**  
Shona WHYTE, Présidente du jury, Professeure  
des Universités, Université Côte d'Azur  
Martin HOWARD, Professeur, University  
College Cork  
Pascale LECLERCQ, Maître de Conférences,  
HDR, Université Paul Valéry Montpellier 3

# L2 French students' comprehension of non-literal language: a replication study on the comprehension of three different types of indirect meaning

*La compréhension du langage non-littéral chez les apprenants de FLE: une étude de réplication sur la compréhension de trois types d'énoncés au sens indirect.*

## **Directrice de thèse**

Amanda EDMONDS, Professeure des Universités, Université Côte d'Azur

## **Présidente du jury**

Shona WHYTE, Professeure des Universités, Université Côte d'Azur

## **Rapporteurs**

Martin HOWARD, Professeur, University College Cork

Pascale LECLERCQ, Maître de Conférences, HDR, Université Paul Valéry Montpellier 3

## Abstract

The ability to understand indirect meaning (utterances for which the literal meaning is different from its intended interpretation) is an important aspect of the pragmatic competence—one that is often neglected in the teaching of foreign languages. Previous research has shown that second language (L2) learners regularly struggle with indirect implicatures or indirect speech acts, and that factors such as proficiency or conventionality of expressions or patterns impact comprehension. One study, published in 2016 by Taguchi, Gomez-Laich and Arrufat-Marques, investigated L2 Spanish learners' comprehension of three types of indirect meaning (indirect refusals, irony, and indirect opinions) which presented different degrees of conventionality or indirectness. Their unexpected findings called into question the idea that conventionality facilitates comprehension. To verify these results, we conducted a close replication of Taguchi et al.'s study with a group of L2 French learners, thus expanding the scope of foreign language comprehension process.

Following Taguchi et al.'s methodology, we used a mixed-method research approach: we assessed our L2 learners' comprehension accuracy and speed through a computer-based multimedia test, and we conducted retrospective interviews with a subset of participants in order to gain some insights into the L2 inferential process. Our quantitative analyses showed significantly lower accuracy and slower response times for the comprehension of certain types of indirect meaning (specifically irony). These results were consistent with Taguchi et al.'s findings. The effect of conventionality of the indirect meaning was not observed: like Taguchi et al., our participants did not have more difficulties with indirect opinions (which were labelled as non-conventional) than they did with indirect refusals (which were considered conventional). Our qualitative analyses point to a significant impact of proficiency: participants who struggled to reach the intended meaning usually admitted to not understanding the target indirect utterances on a basic linguistic level. Nevertheless, participants were able to rely on contextual cues present in our video-assisted listening test, and use them either as supportive inferencing strategies to interpret indirect meaning, or to compensate for their lack of linguistic knowledge or poor listening skills.

## Keywords

Indirect meaning, Replication, L2 French, Comprehension, Pragmatics, Second Language Acquisition

## Résumé

La capacité à comprendre des énoncés au sens indirect (pour lesquels “le dit” et “le signifié” sont différents) est un aspect important de la compétence pragmatique—compétence qui est souvent négligée dans l’enseignement des langues étrangères. Les recherches effectuées précédemment ont montré que les apprenants de langue seconde (L2) éprouvent régulièrement des difficultés face aux implicatures ou aux actes de parole exprimés de façon indirecte, et que des facteurs tels que le niveau de maîtrise de la langue ou la présence (ou l’absence) d’expressions ou de schémas de discours conventionnels affectent la compréhension des apprenants de L2. Une étude, publiée en 2016 par Taguchi, Gomez-Laich et Arrufat-Marques, s’est focalisée sur la compréhension de trois types d’énoncés au sens indirect (le refus indirect, l’ironie et l’expression de l’opinion indirecte) qui représentent des degrés de conventionalité différents. Leurs résultats inattendus ont remis en question l’idée que les expressions et schémas conventionnels facilitent la compréhension. Afin de vérifier ces résultats, nous avons mené une réplique directe de l’étude de Taguchi et al. avec un groupe d’apprenants de français langue étrangère (FLE), permettant ainsi d’élargir le champ de nos connaissances des processus de compréhension dans une autre L2.

En suivant la méthodologie de Taguchi et al., nous avons utilisé une méthode de recherche mixte : nous avons évalué la capacité et la vitesse de compréhension des apprenants de FLE grâce à un test multimédia sur ordinateur, puis nous avons mis en place des entretiens rétrospectifs avec un petit groupe de participants pour essayer d’obtenir des informations sur le processus d’inférence des apprenants lorsqu’ils sont confrontés à des énoncés au sens indirect. Nos analyses quantitatives ont montré une moindre compréhension et des temps de réponses plus lents pour la compréhension de certains types d’énoncés indirects (plus précisément l’ironie). Ces résultats sont en accord avec ceux de Taguchi et collègues. De plus, l’impact du phénomène de conventionalité n’a pas été détecté : tout comme Taguchi et al., nos participants n’ont pas eu plus de difficultés à comprendre les exemples d’opinion indirecte (classifiés comme étant non-conventionnels) que les exemples de refus indirect (considérés comme conventionnels). Nos analyses qualitatives mettent en avant un impact significatif du niveau de langue : les participants qui n’ont pas réussi à comprendre le sens indirect des énoncés ont généralement reconnu ne pas comprendre le sens littéral de ceux-ci à un niveau linguistique de base. Néanmoins, nos participants ont pu s’appuyer sur les éléments contextuels présents dans notre test audio-visuel (gestuelle, expression du visage, intonation, situation, etc.), et ont utilisé

ces repères soit comme sources de stratégies déductives pour interpréter le sens indirect, soit pour compenser leurs lacunes linguistiques et faibles capacités de compréhension auditive.

## **Mots clés**

Non-littéral, Réplication, Français Langue Étrangère, Compréhension, Pragmatique, Acquisition de langue étrangère

## Acknowledgments

Who would have thought when I started this adventure that is the PhD. that, in addition to having to juggle a full-time job, parent two littles, and write a dissertation, I would also have to work through a global pandemic, months of lockdown and online teaching, along with a few medical emergencies? It has been a long journey, but none of this would have been possible without the help and guidance of many people whom I would now like to thank.

First and foremost, I would like to express my deepest gratitude to my *directrice de thèse*, Dr. Amanda Edmonds, without whom this project would not have come to fruition. There were many a time when I was not sure I could do this, but she believed in me, even when I doubted myself. Her support, encouragement, patience and invaluable feedback were absolutely vital to my ability to finish this dissertation.

I would like to thank the members of my dissertation jury, Dr. Howard, Dr. Leclercq, and Dr. Whyte, as well as the professors who served on my *comités de suivi* over the years, for their time, esteemed advice and helpful suggestions.

I am grateful to family members who provided feedback on my instrument during the development phase, and to Célia and Pierre-Francois who graciously agreed to be the actors in my videos. Their help was instrumental in my success and the completion of this project.

Many thanks to friends and colleagues who lent a listening ear when I needed one, cheered me on, helped me brainstorm, or gave me clarity and perspective. Special recognition goes to Dr. Aurore Mroz who was always there for me, mentoring and inspiring me.

Finally, I would like to thank my husband and two children. Ryan took care of the kids on those days I was locked up in my office working on my dissertation. I hope Oliver and Elliot will remember the perseverance and resilience their mom showed throughout this process and that this can be a reminder that even when things are hard and they feel like giving up, they can find the strength to keep going.



## TABLES OF CONTENTS

TABLES OF CONTENTS.....	1
GENERAL INTRODUCTION.....	9
PART I BACKGROUND AND LITERATURE REVIEW.....	13
<b>Background Introduction</b> .....	<b>14</b>
<b>Chapter 1 - Replication</b> .....	<b>15</b>
1.1. Why is replication important?.....	15
1.2. What kind of replications exist?.....	17
1.2.1. <i>Exact replications</i> .....	17
1.2.2. <i>Close replications</i> .....	18
1.2.3. <i>Approximate replications</i> .....	19
1.2.4. <i>Conceptual replications</i> .....	19
1.3. Challenges in replication.....	20
1.3.1. <i>Study design: methodology and research questions</i> .....	20
1.3.2. <i>Result analysis and discussion</i> .....	23
<b>Chapter 2 - Theoretical background</b> .....	<b>26</b>
2.1. Definition: what is indirect meaning?.....	26
2.2. Implicature.....	28
2.3. Cooperative Principle and Conversation Maxims.....	30
2.4. Theories of Politeness (and related principles).....	36
2.5. Relevance theory.....	42
2.6. Indirect meaning processing: The case of irony.....	45
2.6.1. <i>Definition of irony</i> .....	46
2.6.2. <i>A literature review of the role of literal meaning in the mechanisms of irony production and comprehension in L1</i> .....	48
2.6.3. <i>Functions of verbal irony</i> .....	52
2.6.4. <i>Accessing the ironic interpretation</i> .....	54
2.7. Indirect meaning from the perspective of Conversation Analysis.....	62
2.7.1. <i>Sequence organization and turn-taking</i> .....	63
2.7.2. <i>Preference organization</i> .....	64
<b>Background conclusion</b> .....	<b>66</b>
<b>Chapter 3 - L2 Indirect meaning comprehension: a literature review</b> .....	<b>68</b>

3.1.	Impact of proficiency on L2 comprehension .....	75
3.2.	Impact of indirect meaning types on L2 comprehension.....	85
3.3.	Inferencing Strategies used for L2 comprehension .....	96
<b>Chapter 4 - A close look at Taguchi et al.'s (2016) study.....</b>		<b>105</b>
4.1.	Categories of indirect meaning .....	106
4.1.1.	<i>Indirect refusals</i> .....	106
4.1.2.	<i>Indirect opinions</i> .....	107
4.1.3.	<i>Irony</i> .....	108
4.1.4.	<i>Potential challenges with the three categories</i> .....	110
4.2.	Instrument development.....	111
4.3.	Experiment.....	117
4.3.1.	<i>Test Participants</i> .....	117
4.3.2.	<i>Survey administration</i> .....	118
4.3.3.	<i>Retrospective interviews</i> .....	119
4.4.	Results and discussion .....	122
4.4.1.	<i>Effects of indirect meaning types on comprehension</i> .....	123
4.4.2.	<i>Posthoc Analyses of Retrospective Verbal Interviews</i> .....	126
4.5.	Limitations and potential for replication and extension .....	130
<b>Part I - Conclusion.....</b>		<b>133</b>
<b>PART II EXPERIMENTAL STUDY .....</b>		<b>135</b>
<b>Introduction .....</b>		<b>136</b>
<b>Chapter 5 - Instrument development in the context of a replication study .....</b>		<b>137</b>
5.1.	Multimedia Listening Test Development .....	137
5.1.1.	<i>Non-literal dialogues</i> .....	138
5.1.2.	<i>Translation choices and modifications</i> .....	139
5.1.3.	<i>Fillers</i> .....	146
5.1.4.	<i>Feedback and final modifications</i> .....	148
5.1.5.	<i>Videos</i> .....	154
5.2.	Instrument development.....	156
5.2.1.	<i>Multiple-choice questionnaire tasks</i> .....	157
5.2.2.	<i>Online survey questionnaire</i> .....	160
5.2.2.1.	Presentation of the items .....	160
5.2.2.2.	Challenges and solutions.....	161
<b>Chapter 6 - Experimental study - Pilot with native speakers.....</b>		<b>164</b>
6.1.	Methods.....	164

6.1.1.	<i>Participants</i> .....	164
6.1.2.	<i>Procedures</i> .....	165
6.2.	Native speaker performance and impact of indirect meaning on comprehension	166
6.2.1.	<i>Data Analysis</i> .....	166
6.2.2.	<i>Results</i> .....	168
6.2.2.1.	Comprehension accuracy .....	168
6.2.2.2.	Comprehension speed .....	169
6.2.3.	<i>Discussion</i> .....	171
6.3.	Results and selection of items to be removed .....	174
6.3.1.	<i>Data analysis</i> .....	174
6.3.2.	<i>Results</i> .....	174
6.3.3.	<i>Identification of items to be removed</i> .....	177
<b>Chapter 7 - Experimental study – L2 learners</b> .....		<b>180</b>
7.1.	Methods.....	180
7.1.1.	<i>Participants in the L2 study</i> .....	180
7.1.2.	<i>Learners’ self-assessment of language proficiency</i> .....	181
7.1.3.	<i>Procedures</i> .....	183
7.1.3.1.	IRB.....	183
7.1.3.2.	Survey administration .....	184
7.1.4.	<i>Data analysis</i> .....	186
7.2.	Results.....	187
7.2.1.	<i>Comprehension accuracy</i> .....	187
7.2.2.	<i>Comprehension speed</i> .....	189
7.3.	Discussion .....	192
7.3.1.	<i>Answers to our research questions</i> .....	192
7.3.2.	<i>Conventionality effect</i> .....	194
7.3.3.	<i>Comparison with Taguchi et al. (2016)</i> .....	196
7.3.4.	<i>Reprocessing of the native speaker dataset</i> .....	198
7.3.4.1.	Data analysis .....	198
7.3.4.2.	Results.....	199
<b>Part II - Conclusion</b> .....		<b>204</b>
<b>PART III RETROSPECTIVE VERBAL INTERVIEWS</b> .....		<b>205</b>
<b>Introduction</b> .....		<b>206</b>
<b>Chapter 8 - Methods</b> .....		<b>207</b>
8.1.	Participants.....	207
8.2.	Item selection .....	208

8.3.	Procedure .....	210
8.4.	Coding choices for the replication study .....	211
8.4.1.	<i>Visual and auditory cue strategies</i> .....	215
8.4.2.	<i>Background knowledge strategy</i> .....	217
8.4.3.	<i>Keyword strategies</i> .....	218
8.4.4.	<i>Logical reasoning strategies</i> .....	220
8.4.5.	<i>Additional Strategies</i> .....	220
8.4.5.1.	Setting/situation cues .....	221
8.4.5.2.	Familiarity with the characters.....	221
8.4.5.3.	Adjacency pair rule .....	222
8.4.5.4.	Process of elimination.....	223
8.5.	Data analysis .....	223
<b>Chapter 9 - Analyses of Retrospective Verbal Interviews: Results and Discussion...</b>		<b>226</b>
9.1.	Inferencing strategy use and effectiveness among high- and low-scoring participants.....	226
9.1.1.	<i>Access to intended meaning vs. access to literal meaning of the target utterance</i> .....	226
9.1.2.	<i>Differences in strategy use by participant group</i> .....	229
9.1.2.1.	Keyword inferencing.....	230
9.1.2.2.	Logical reasoning .....	234
9.1.2.3.	Visual cues .....	236
9.1.2.4.	Process of elimination.....	239
9.1.2.5.	Auditory cues .....	241
9.2.	Inferencing strategies as a function of item type .....	245
9.2.1.	<i>Indirect refusals</i> .....	247
9.2.2.	<i>Irony</i> .....	249
9.2.2.1.	Focus on IRO 8: the issue with Harry Potter .....	255
9.2.3.	<i>Indirect opinions</i> .....	259
<b>GENERAL CONCLUSION</b> .....		<b>262</b>
<b>Summary of findings and importance of replication</b> .....		<b>263</b>
<b>Pedagogical implications</b> .....		<b>267</b>
<b>Limitations and suggestions for future research</b> .....		<b>268</b>
<b>REFERENCES</b> .....		<b>271</b>
<b>APPENDICES</b> .....		<b>289</b>
<b>Appendix A. Summary of Taguchi's studies on L2 implicature comprehension</b> .....		<b>289</b>

<b>Appendix B. Sample Comparison of Original and Modified Text in Multimedia Listening Test for Indirect Refusal, from Taguchi et al. (2013) – Appendix – p. 698 291</b>	
<b>Appendix C. Test items (French and English version).....</b>	<b>292</b>
<b>Appendix D. Dialogue, video and MCQ lengths .....</b>	<b>334</b>
<b>Appendix E. Opinion/judgement survey .....</b>	<b>337</b>
<b>Appendix F. Reviewers dialogue ratings .....</b>	<b>339</b>
<b>Appendix G. Biographical data collection.....</b>	<b>341</b>
<b>Appendix H. Test instructions (L2 – English version) .....</b>	<b>342</b>
<b>Appendix I. Pilot Study recruitment message.....</b>	<b>343</b>
<b>Appendix J. ACTFL Can-Do Statements.....</b>	<b>344</b>
<b>Appendix K. Informed Consent Form.....</b>	<b>345</b>
<b>Appendix L. Learner Accuracy Random Effects .....</b>	<b>348</b>
<b>Appendix M. Descriptive Statistics for the learners’ response times for the full dataset (including correct and incorrect answers).....</b>	<b>351</b>
<b>Appendix N. Learner response time Accuracy Model - Random effects .....</b>	<b>352</b>
<b>Appendix O. Difference between high and low-scoring mean for follow-up interview item selection .....</b>	<b>355</b>

## List of Tables

Table 1 <i>Reviewed L2 indirect meaning studies (5 pages total)</i> .....	69
Table 2 <i>Taguchi et al. (2016) – L2 Spanish Learners’ Comprehension Accuracy Scores</i> ....	122
Table 3 <i>Taguchi et al. (2016) – L2 Spanish Learners’ Comprehension Response Times</i> .....	123
Table 4 <i>Taguchi et al. (2016) – Frequency of Inferencing Strategies by Participant Group</i>	126
Table 5 <i>Taguchi et al. (2016) – Frequency of Inferencing Strategies by Item Type</i> .....	128
Table 6 <i>Word count for the French Dialogue</i> .....	149
Table 7 <i>Length of videos (in seconds)</i> .....	149
Table 8 <i>MCQ word count (English and French)</i> .....	160
Table 9 <i>Native Speakers’ Accuracy Scores</i> .....	168
Table 10 <i>NSs’ Response Times on correct answers</i> .....	170
Table 11 <i>Item Accuracy Scores</i> .....	175
Table 12 <i>Items Response Times</i> .....	176
Table 13 <i>Learners’ Accuracy Scores</i> .....	188
Table 14 <i>Accuracy Model results</i> .....	189
Table 15 <i>Learners’ Response Times on correct answers</i> .....	190
Table 16 <i>Response Time Model results</i> .....	191
Table 17 <i>Comparison of Learners’ Accuracy Scores between Taguchi et al. and our data</i> .	196
Table 18 <i>NSs’ Accuracy Scores (retained items only)</i> .....	200
Table 19 <i>NSs’ Response Times on correct answers (retained items only)</i> .....	201
Table 20 <i>Participants’ accuracy scores</i> .....	208
Table 21 <i>Accuracy rate on interview items per group</i> .....	227
Table 22 <i>Frequency of Inferencing Strategies by Participant Group</i> .....	229
Table 23 <i>Frequency of Inferencing Strategies by Item Categories</i> .....	246
Table 24 <i>Frequency of Inferencing Strategies by Item Categories based on group</i> .....	246
Table 25 <i>Summary of Taguchi’s studies on L2 implicature comprehension</i> .....	289
Table 26 <i>Dialogue, video and MCA lengths</i> .....	334
Table 27 <i>Reviewers dialogue ratings</i> .....	339
Table 28 <i>Learner Accuracy Model - Random Intercept</i> .....	348
Table 29 <i>Learner Accuracy Model - Random Slope</i> .....	348
Table 30 <i>Learners’ Response Times</i> .....	351
Table 31 <i>Response Time Model random effects - Participant Intercept</i> .....	352
Table 32 <i>Response Time Model random effects - Participant/Condition Intercept</i> .....	352

Table 33 *Difference between high and low-scoring mean for follow-up interview item selection*..... 355

## List of Figures

Figure 1 <i>Taguchi et al. (2016) – IRF1</i> .....	114
Figure 2 <i>Taguchi et al. (2016) – OPI6</i> .....	115
Figure 3 <i>Taguchi et al. (2016) – IRF9</i> .....	116
Figure 4 <i>Taguchi et al. (2016) – Categories of Inferencing Strategies and Sample Interview Excerpts</i> .....	120
Figure 5 <i>IRO4 – task comparison between Taguchi et al. (2016) and our version</i> .....	145
Figure 6 <i>IRO6 - dialogue</i> .....	147
Figure 7 <i>Filler 2 and Filler 3 - dialogues</i> .....	148
Figure 8 <i>IFR1 – task comparison between Taguchi et al. (2016) and our version</i> .....	158
Figure 9 <i>NSs' Accuracy Scores – bar graph with error bars</i> .....	169
Figure 10 <i>NSs' Response Times on correct answers - bar graph with error bars</i> .....	170
Figure 11 <i>Items Accuracy Scores – grouped boxplot</i> .....	175
Figure 12 <i>Items Response Times – grouped boxplot</i> .....	177
Figure 13 <i>Learners' Accuracy Scores - bar graph with error bars</i> .....	188
Figure 14 <i>Learners' Response Times on correct answers - bar graph with error bars</i> .....	191
Figure 15 <i>NSs' Accuracy Scores (retained items only) - bar graph with error bars</i> .....	200
Figure 16 <i>NSs Response Times on correct answers (retained items only)</i> .....	202
Figure 17 <i>Categories of Inferencing Strategies and Sample Interview Excerpts from Taguchi et al. (2016)</i> .....	212
Figure 18 <i>Categories of Inferencing Strategies with example from our interviews</i> .....	214



# **GENERAL INTRODUCTION**

The ability to understand implied speaker intention is an important part of communication which falls under the pragmatic competence. Taguchi (2007) defines the pragmatic competence as “the ability to produce meaning in a socially appropriate manner and to interpret meaning, explicitly or implicitly stated, according to contexts” (p. 313-314). While researchers emphasize the importance of developing pragmatic skills as part of the second language (L2) learning process, studies show that L2 learners often struggle with indirect meaning comprehension. Through our own experience as an educator, we have observed that L2 French learners sometimes completely disregard the intended indirect interpretation of an utterance to only focus on its literal meaning, leading to misunderstanding and communication breakdown. This is particularly true with regard to sarcasm and irony, as our American students often note that French people have a very particular sense of humor that can be challenging for American speakers.

Many studies within the field of L2 pragmatics have investigated L2 comprehension of different types of indirect meaning, such as indirect refusals, indirect apologies, indirect requests or indirect opinions, though predominantly in L2 English. Results usually point to an impact of proficiency as well as the facilitating effect of conventionality of indirect meaning (either in the form of conventional linguistically-encoded expressions, or through conventional structures or patterns of discourse). Indeed, more proficient L2 learners were generally able to understand indirect meaning better and faster than less proficient learners, and more conventional types of indirect meaning were easier to comprehend than less conventional ones. One study however, published in 2016 by Taguchi, Gomez-Laich and Arrufat-Marques, seemed to call this conventionality effect into question. They set out to investigate L2 Spanish comprehension of three different types of indirect meaning varying in degrees of conventionality: indirect refusals, irony, and indirect opinions. To that effect, they used a novel multimedia listening test that offered a unique combination of visual and oral input. Given their unexpected findings and our personal interest in the exploration of pragmatic inferential mechanisms (and our students’ expressed struggle with irony), we decided to conduct a close replication of Taguchi et al.’s study to see if we could verify and generalize their findings with L2 French learners.

Undertaking this replication seemed particularly important to us, given the limited number of replications published in the field of Applied Linguistics, and the fact that, in recent years,

authors (Markee, 2017; Marsden, Morgan-Short, Thompson and Abugaber, 2018a; Marsden, Morgan-Short, Trofimovich, Ellis, 2018b; Porte & McManus, 2019; Porte & Richards, 2012) have highlighted the need for replication in order to confirm and validate unexpected or influential findings, especially when the latter may be used to inform our pedagogical practices. Our dissertation therefore aims to respond to the call for replication studies in second language acquisition research, and expand the scope of foreign language indirect comprehension by using L2 French as its target language.

Following Taguchi et al.'s study, our project seeks to answer the following research questions to investigate comprehension of indirect meaning by L2 French learners through a multimedia listening task:

- 1) To what extent are L2 French learners able to understand indirect meaning?
- 2) How do different types of indirect meaning impact L2 French learners' comprehension accuracy and speed?
- 3) What inferential strategies do L2 French learners use and how successful are they?

This dissertation is divided into three main parts. In [Part I](#), we will first discuss the value of replication in Applied Linguistics: why is replication important in our field, what are the different types of replications that researchers can undertake, and what challenges they may face? ([Chapter 1](#)). We will then present our theoretical frameworks starting with a definition of indirect meaning, followed by an overview of various theories that shed light on the nature of indirect meaning and the ways people process it ([Chapter 2](#)). In [Chapter 3](#), we will review existing literature investigating L2 indirect meaning comprehension, focusing specifically on three factors that have been found to impact comprehension: proficiency, indirect meaning type and inferential strategy use. Finally, [Chapter 4](#) will be dedicated to a review and critical analysis of the study we are replicating (Taguchi et al., 2016), and will include a discussion on its potential methodological issues and limitations.

[Part II](#) will focus on the quantitative part of our experimental study with the objective of answering our first two research questions. In [Chapter 5](#), we will discuss the development of our multimedia instrument (from the creation of the task dialogues, to the production of our video recordings, and the design of our online survey). [Chapter 6](#) will be dedicated to the pilot study with native speakers: we will present our methods and results, and briefly discuss our findings as they compare to Taguchi et al.'s. We will also explain how the pilot study results

were used to refine our instrument and eliminate some items for the final test used with our L2 learners. The L2 learner test will be the focus of [Chapter 7](#). We will present our methods, results with regard to L2 learners' comprehension accuracy and speed, and answers to our research questions. We will also discuss our L2 French learners' performance as it compares to Taguchi et al.'s L2 Spanish learners, as well as to our native speakers, in order to reveal potential similarities and differences in the context of our replication.

[Part III](#) will concentrate on the retrospective interviews we conducted with a subset of participants in order to gain some insights into L2 learners' inferential process. After presenting our methods, we will discuss our coding choices, and establish a list of inferencing strategies we identified in our data ([Chapter 8](#)). In [Chapter 9](#), we will analyze and discuss our qualitative results to reveal differences among the two groups of interview participants (low and high proficiency) as well as different patterns of strategy use as a function of indirect meaning types.

Finally, our general conclusion will summarize our major findings and the contributions of our replication to the field of L2 Pragmatics. We will offer some pedagogical recommendations, and expose some of our study limitations along with some suggestions for improvements and ideas for future research.

# **PART I**

## **BACKGROUND AND LITERATURE REVIEW**

## Background Introduction

This dissertation aims to study L2 comprehension of indirect meaning, specifically, how L2 French learners understand (or not) indirect refusals, indirect opinions and irony. Production and comprehension of non-literal utterances have been studied through the lens of various related fields—psycholinguistics, neurolinguistics, sociolinguistics, etc.—all the while focusing on an array of documents: from literary written texts (e.g., poetry, fiction, etc.), to oral interaction (recordings of spontaneous exchanges between friends). The use of indirect meaning belongs to the pragmatic competence, defined by Taguchi (2005, p. 543) as “the ability to comprehend and produce meaning in context.” This particular aspect has generated many theories, in particular in L1 (from Grice’s implicatures and Conversational Maxims to the Relevance Theory).

The first part of our dissertation will therefore present an overview of the theories relating to indirect meaning in L1, as well as a more focused review of L2 research studies on the same topic. Before that however, we will devote our first chapter ([Chapter 1](#)) to the concept of replication (what it means, why it is important to carry out replication in L2 research, and what some potential challenges might be). [Chapter 2](#) will present the theoretical background for indirect meaning processing in general (as it relates to L1). In [Chapter 3](#), we will review the relevant L2 research studies that focus on indirect meaning. Finally, we will dedicate the last chapter of this first part ([Chapter 4](#)) to a review and analysis of the research study this dissertation is replicating: Taguchi, Gomez-Laich and Arrufat-Marques’s paper on the Comprehension of Indirect Meaning in Spanish as a Foreign Language (2016).

## Chapter 1 - Replication

This first chapter will start by explaining why replication is important and needed in the field of applied linguistics. Then we will present an overview of the different types of replications. Finally, we will discuss some of the challenges that researchers may face when tackling replication studies.

### 1.1. Why is replication important?

In recent years, replication research has been gaining in popularity and importance in the field of Applied Linguistics (McManus, 2021). However, several authors (Markee, 2017; Marsden, Morgan-Short, Thompson and Abugaber, 2018a; Marsden, Morgan-Short, Trofimovich, Ellis, 2018b; Porte, 2012; Porte & McManus, 2019; Porte & Richards, 2012) point out the critically low number of replications published each year: Marsden et al. (2018a) estimate that only 1 in 400 articles published in L2 research journals was a self-labelled replication. Reasons that may explain this lack of replication include a perceived risk of not being published (unless initial findings are reproduced), poor availability of materials and raw data (which we will discuss in [section 1.3](#)), and an enduring notion that replication has a low impact and lack of prestige. Yet, the concept of replication is one that has long been accepted in other scientific fields: the “need to verify or test previous hypotheses and further probe research outcomes [...] as a useful means of contributing to the body of knowledge and regarded as an essential and established part of the research process” (Porte & McManus, 2019, p. 4). As this excerpt from the editorials of *Nature*, Vol. 442 (2006) explains, in many fields, replication is already an integral part of the research validation process (p. 330):

In some disciplines, fresh work that builds on a published result will involve fully replicating the work that led to that result. In others, including many branches of biology, scientists see the process as validation, rather than replication. The expectation isn't necessarily that another scientist will reproduce exactly what they did, but that the data and methodology are strong enough to withstand inferences and have new experiments built upon them.

Indeed, the purpose of replication is to test the reliability of a previous study's findings by validating (or providing evidence against) its results, or to generalize findings in the same or under different conditions (Marsden et al., 2018a; Porte & McManus, 2019). One may want to replicate a study because its findings are very influential in the field but have not been replicated

before, or because the initial<sup>1</sup> study challenges existing theories or yields unexpected results (Marsden et al., 2018a). Another reason may come from the need to validate results because of concerns with research design or practices (e.g., surprising results, unreported effect sizes, low statistical power, etc.—cf. Marsden et al, 2018a; Porte & McManus; 2019; Porte & Richards, 2012). As Porte and Richards (2012) state, by doing so “we are also contributing to the research community by revisiting earlier assumptions or conclusions, confirming or otherwise what has previously been produced or thought” (p. 285). The importance of replication is particularly true in the field of applied linguistics where “much of our theoretical understanding is built on a small number of influential studies that have not been replicated” (McManus, 2021, p. 3), and where findings may have impactful consequences on pedagogical practices:

The consequent absence of replication studies is especially problematic if we accept that error, both of a methodological and conceptual kind, inevitably accompanies much of the research we do. Such unattended error is particularly disquieting in a field with the kind of practical objectives we espouse and in which we might reasonably expect research to regularly and reliably inform educational policy on L2 teaching. It is also reasonable to imagine that basing key pedagogical decisions on the outcomes of a series of studies with undetermined and/or unverified amounts of error may even have proved counterproductive in some learning contexts in the past (Porte, 2012).

Porte & Richards, 2012, p. 285

As a result, several authors (Marsden et al., 2018, Markee, 2017; Porte & McManus, 2019; Porte & Richards, 2012) insist on the need to work collectively on a replication research agenda that will “promote knowledge transfer and construction across that research rather than mere knowledge accumulation” and “encourage the kind of healthy curiosity and skepticism towards research outcomes that will help move us forwards – rather than sideways” (Porte & Richards, 2012, p. 291). Because replications aim to validate and generalize findings, some authors view it as the “gold standard” of research evidence (Jansy, Chin, Chong & Vignieri, 2011, p. 1225) and a “linchpin of the scientific process” (Let’s replicate, 2006, p. 33).

---

<sup>1</sup> The use of the term “initial” study instead of “original” study is recommended by Marsden et al. (2018a) to refer to a study that has been replicated because “studies are rarely, if ever truly original in the sense of being a completely novel idea. Also, original carries negative connotations for its replication because it could imply that anything that is not original cannot share characteristics broadly associated with originality” (p. 326).



## 1.2. What kind of replications exist?

Most researchers agree in the categorization of replication into three or four different types, though terminology and definitions sometimes vary slightly. Porte (2012) and Marsden et al. (2018a) recommend a classification in three types: exact or literal replication; partial, approximate or systematic replication; and conceptual or constructive replication. However, as Marsden et al. (2018a) explain, “the extent to which this recommendation has been adopted by the field in a systematic manner remains unclear” (p. 326). For the purpose of this dissertation, we will adopt the nomenclature presented in Porte and McManus (2019) (which is also the same as McManus, 2021): replications are categorized as *exact*, *close*, *approximate* or *conceptual*. Our next sub-sections will attempt to define each type of replication.

### 1.2.1. Exact replications

As one can guess, an exact (or direct, literal, strict or virtual) replication is one that follows the initial study procedure without any modification. The researcher will make it a point to keep a similar data sample (same participant characteristics), the same research questions, and the same design, coding and analyses (McManus, 2021). Marsden et al. (2018a) reference Lykken (1968) who illustrates a literal replication as one that would require recruiting additional participants for the same study and thus repeating the same research procedure with a similar sample. Most linguists, however, agree that carrying out exact replication in the social sciences (and, thus, in applied linguistics) would be very challenging, if not impossible due to the fact that “there are always differences between the initial study and the replication” (McManus, 2021, p. 6) and that any changes in conditions, context, participants (unless one re-uses the same exact subjects) would introduce a modification that would disqualify the study as an exact replication. This reason has often been given as a justification for the lack of replication in social sciences, and Porte and Richards (2012) quote LeCompte and Goetz (1982) objecting to the possibility of doing replication in qualitative ethnography due to the “personalistic” nature of the field and the fact that if “human behavior is never static, no study can be replicated exactly, regardless of the methods and designs employed” (Le Compte & Goetz as cited in Porte & Richards, 2012, p. 288). This likely explains why Marsden et al. (2018a) reported a very low number of exact replications in the social sciences in their paper focusing on a systematic review of replication in L2 research. Yet, as Porte and Richards (2012) point out, “[r]eplication is not often undertaken, as it is understood to be limited to exact replication studies which exist

to confirm the outcomes of the original studies” (p. 284), and considering replication from this very narrow definition is a mistake that Porte and Richards have been working to correct. Indeed, replications do not have to be literal to be useful and valuable to the field, as we will see with the next three categories of replication.

### *1.2.2. Close replications*

A close replication is one that follows the initial study research design and procedures closely but that introduces one modification. This means that “[a] close replication study could modify, for example, the instruction, or the outcome measure, or the L2, or the L1 in order to understand how modifying that variable influenced the study’s outcomes” (McManus, 2021, p. 6). This definition of a close replication is one that McManus (2021) and Porte and McManus (2019) follow, but it should be noted that other authors (Abbuhl, 2012; Markee, 2017; Marsden et al., 2018a, Marsden et al., 2018b; Porte, 2012; Porte & Richards, 2012) use the term partial or approximate replication to encompass close replication<sup>2</sup>: “one or two non-major variables are changed to allow for eventual comparability between both studies and their outcomes” (Porte & Richards, 2012, p. 286).

In Applied Linguistics, close replications often call for testing new or different conditions, L1 background, target language, proficiency, amount of instruction, etc. (Marsden et al., 2018a). As Porte and McManus (2019) state (p. 73):

One of the principal reasons for conducting a replication is to increase the confirmatory power of the original study. Replications proceed by degrees: with every pertinent modification in subsequent replications the confirmatory power of the study may increase and, potentially, be generalized to further or wider applications.

By manipulating one variable, we have a better chance at understanding how it may impact the particular phenomenon under study, and close replications play a crucial role in fine-tuning our findings to make them more generalized: “A close replication is perhaps the clearest way to advance knowledge and understanding in a field because it sets out to investigate how intentionally modifying a single variable impacts the study’s findings” (McManus, 2021, p. 6).

---

<sup>2</sup> Additionally, Markee’s (2017) example of an exact replication is one that would be labelled as close replication by Porte and McManus (2019). Indeed, she illustrates the concept of exact replication using a study by Moerman (1997) who re-produces Schegloff, Jefferson and Sacks’s (1977) study on Conversation Analysis (speaker preference for self-repair) in which he changed the target language (from English to Thai).

Of course, researchers doing close replications will have to be very careful to control all other aspects of the replication to ensure that only one variable is being modified. We will discuss the challenges that this presents in [section 1.3](#).

### *1.2.3. Approximate replications*

As mentioned in the previous sub-section, some authors include close replications in the approximate or partial replication category (or use the term “partial” replication to refer to “a study that intentionally changes only one significant component of the initial study,” Marsden et al., 2018a). McManus (2021) and Porte and McManus (2019) however, separate these two types of replications in the sense that close replication should only see the modification of one single variable, whereas approximate or partial replications are “similar to a close replication but with more room for modification” (McManus, 2021, p. 7). Porte and McManus (2019) see approximate replications as a step that comes after a series of close replications, and as a result will include two or more modified variables (Porte & McManus, 2019). Approximate replications therefore seek to compare findings with that of the initial study in light of the previous close replications and the reported effects of the modified variables (Porte & McManus, 2019, p. 78).

### *1.2.4. Conceptual replications*

Conceptual (or constructive) replications are further away from the initial study: while they will focus on similar research questions or hypotheses as the initial study, different methodology might be used (e.g., different design, sample, data collection procedures, etc.). Porte and McManus (2019, p. 83-84) explain that with a conceptual replication, researchers will:

[A]ttempt to widen the application, relevance, or generalizability of the underlying theory or hypothesis of the original study by, for example, using a completely different database and related collection procedures (e.g., observation against self-report, qualitative supplementing original quantitative data), operational definition or methodology, or method of analysis.

Indeed, conceptual replications do not aim for a direct comparison with the initial study. As such, they cannot claim to test the validity or reliability of the initial study itself (Porte & McManus, 2019, p. 91). However, if findings from the conceptual replication(s) of an initial study support the initial theory, we can have confidence in the application and generalizability

of the initial findings since the initial study's outcomes "were not just artefacts of the original methodology" (Porte & Richards, 2012, p. 286). Yet, due to the wider range of modifications present in conceptual replication, and the fact that it aims to extend or build upon the initial study's theory(ies), Porte and McManus (2019) explain that some researchers question the use of the term replication for these types of studies that may be seen as mere extension studies by some: "how sure can we be that the original study and the conceptual replication of it are actually observing the same phenomenon – remember you might well have chosen to provide a different definition of the main construct used" (p. 93). Still, according to them, "well-planned and well-argued conceptual replications" (p. 94) that support findings from the original study should be a clear indication to "sit up and listen, for something is clearly emerging as a result of work" (p. 94).

### 1.3. Challenges in replication

As we mentioned in the previous section, exact, literal or direct replications are often deemed "an unachievable objective" (Porte & McManus, 2019, p. 72), due their nature and the need to follow the same exact methodology as the replicated study. Challenges can also arise when attempting close or partial replications. We will now discuss the various issues researchers may encounter when attempting a replication, as well as considerations or recommendations authors provide to address these challenges.

#### *1.3.1. Study design: methodology and research questions*

As stated before, the purpose of a replication is to carry out a study that repeats most or some of the elements of an initial study in order to validate, confirm (or disprove), or generalize its findings. However, reproducibility (whether a replication generates similar findings as the initial study) is not necessarily the end goal: Discrepancies are valuable as well and should encourage further investigation: "Contradictory outcomes are just as interesting as consistent ones and, it must be emphasized, just as worthy of dissemination" (Porte & Richards, 2012, p. 286). Nevertheless, the necessary condition to be able to compare the replication and initial study is adherence to "rigorous standards in terms of how data collection procedures are developed and represented" (Porte & Richards, 2012, p. 288). Indeed, when conducting close or partial replications, researchers must be very careful to only modify the variable(s) on which

they are planning to focus, and follow the rest of the initial study methodology as closely as possible. Any other minor modifications that are deemed unsubstantial should still be discussed for their potential impact on the replication results. One of the potential challenges in replicating all the aspects of the initial study (sample, data collection context, instrument, etc.), is the lack of details provided in the initial study. In their book *Doing Replication Research in Applied Linguistics*, Porte and McManus (2019) illustrate the process of evaluating and critiquing an initial study with the goal of replicating it, using Bitchener and Knoch's (2010) paper on written corrective feedback (CF) as their "source study" (they use this study as an example for various phases of the replication process from planning and feasibility, to replication decision making, carrying out and write up). Upon examining the methodology Bitchener and Knoch used in their initial study, Porte and McManus (2019) acknowledge that very little information about the recruitment of participants was provided in the published article. Porte and McManus explain that recruitment criteria would have been helpful to account for "possible attrition and exclusion rates" (p. 107). Furthermore, while an adequate amount of details is shared regarding participants (number, age, background, L2, class), some information was not reported. For example, the paper did not say which topics were covered in the academic writing course the participants were enrolled in, or give the exact background of the participants—the study only indicates that "Most of the participants were from a range of East and South Asia countries" (Bitchener & Knoch, 2010, as cited in Porte & McManus, 2019, p. 107).

In addition to unreported methodological details, availability of materials (including data collection instrument) and accessibility to raw data are two issues that many authors consider as a main obstacle for the robustness of replications. Marsden et al. (2018b) reference Derrick (2016) who reported a very low availability of research materials in published articles (only 17% of the materials were available in the articles or on online platforms) and Marsden and colleagues (2018b) found that about 27% of the study materials were readily available in their sample of L2 research articles. The consequence of this poor availability is that "future researchers wishing to systematically extend prior studies must either recreate materials, thus introducing unplanned heterogeneity, or work directly with the initial study's authors, thus introducing potential bias" (Marsden et al., 2018b, p. 312). Likewise, not being able to access initial studies' raw data can be problematic as it "prevents replication researchers from validating previous analyses to ascertain the reproducibility of the findings and blocks researchers from combining their own data with previous data sets" (Marsden et al., 2018b, p. 311). Porte and McManus (2019) suggest writing the authors of the initial study if materials

and data are not accessible through supplements or online platforms.

To increase transparency of research, authors (Marsden et al., 2018a; McManus, 2021; Porte & Richards, 2012, to only reference a few) encourage the development of open science research practices that would boost replication feasibility, in particular the use of online depository like IRIS (where authors can publicly and freely share their materials, data collection instrument, and raw data), preregistration in registered reports (a way for researchers to plan and organize future research and receive potential feedback before they begin collection and analyses phases), and the use of open access and preprints so that future researchers can easily (and freely) access articles.

Whether it is because the instrument had to be re-created based on incomplete description in the initial study, or because some of the procedural details (instructions—which Porte & McManus identified as one area that is often omitted in articles—tasks, steps, etc.) were missing or unclear, it will be important for researchers attempting a replication to provide additional information in their write-up to account for potential methodological differences that may explain the presence of different findings. The additional details will also be useful for potential follow-up replications. Therefore, highlighting both similarities and discrepancies in methodologies between the initial study and the replication is of the utmost importance.

When it comes to research questions, we would expect the replication study to follow similar research questions as the initial study since the purpose of the replication is to investigate or further refine a theory or hypotheses presented in the initial study. Some elements of the research questions may be different of course. Porte and McManus (2019) illustrate this point with a potential replication of Bitchener and Knoch (2010) in which they modify the proficiency level of the participants. Bitchener and Knoch's (2010) study focused on advanced learners and their research questions were as follow (p. 211):

1. Does advanced learner accuracy in the use of two functions of the English article system improve over a 10-week period as a result of written CF?
2. Does advanced learner accuracy in the use of two functions of the English article system vary according to the type of written CF provided?

Porte and McManus (2019)'s proposed replication on the other hand, would focus on upper-intermediate level participants. They therefore propose a first set of research questions that

mirror those of Bitchener and Knoch (2010) but use “upper-intermediate” instead of “advanced” to refer to the participant group (p. 103):

1. Does upper-intermediate learner accuracy in the use of two functions of the English article system improve over a 10-week period as a result of written CF?
2. Does upper-intermediate learner accuracy in the use of two functions of the English article system vary according to the type of written CF provided?

Porte and McManus also suggest “include[ing] additional information to improve the clarity and precision of research questions” (p. 103) such as amount of time between post-test and delayed post-test, and precisions regarding the two types of articles on which the study focuses. They propose a second set of research questions with the addition of the expression “to what extent” instead of “does,” which opens the door to an inquiry into the “degrees of improvement instead of the binary questions of “improvement” versus “no improvement”” (p. 103):

- 1a. To what extent does providing written CF improve the accuracy of first and anaphoric mentions in written L2 English immediately after instruction (at post-test) and eight weeks later (at delayed post-test)?
- 2a. To what extent do different types of written CF (direct, indirect, direct + oral review) improve the accuracy of first and anaphoric mentions in written L2 English?

All the above challenges are things that researchers must keep in mind when planning a replication. In the next sub-section, we will review potential challenges or considerations that may come up during the analytical and discussion phases of the replication.

### *1.3.2. Result analysis and discussion*

In order to replicate an experimental study, enough details about its coding and data analysis must be provided in the initial study for researchers to evaluate and interpret its findings. One common issue that several authors (Marsden et al., 2018a; Plonsky, 2015; Porte & McManus, 2019) discuss is the use of  $p$ -value as reaching statistical significance at  $p < 0.05$ , without reporting any effect sizes. Porte and McManus (2019) indicate that even when the effect size is not reported in the initial study, researchers may be able to calculate it based on other provided statistics and it will be important to do so in order to produce a comparative analysis with the findings of the replication. Therefore, following the initial study’s analytical process is not enough: researchers will want to improve on the statistical analyses for increased validity and

confidence in the findings:

[W]e must also be aware of the impact of particular analytical decisions on outcomes, and so we will also consider alternative procedures, if appropriate, which might increase the transparency and robustness of the analysis. When it comes to writing up our replication study's analysis we must also ensure that we fully document any changes followed by justifications for them.

Porte & McManus, 2019, p. 121.

The type of statistical tests used to analyze the data collected in the replication study may not be the same as the tests used in the initial study as statistical analyses usually depend on the data distribution. Careful consideration should therefore be given to check test assumptions and normality (a step that can be omitted or unreported in studies—an issue that the replication should address). Furthermore, if different ranges were used to calculate or interpret effect sizes, researchers may want to present two sets of results: one that will be calculated according to the initial study's procedure, so that a comparison between the initial study and the replication's findings can be made, and one set of effect sizes that may be using different (but maybe more accurate) benchmarks (following McManus & Marsden's (2018) partial replication).

Another consideration that can impact the interpretation of the results will be whether researchers will want to change the operationalized definitions of the variables in the replication. Drawing from their analysis of their source study (Bitchener & Knoch, 2010), Porte and McManus (2019) comment on the fact that Bitchener and Knoch operationalized accuracy of the use of the English articles “a” and “the” as a percentage of correct usage for all occasions. This poses two issues according to Porte and McManus: first both articles are combined for the analysis, and even though they are presented as “two different functions with correspondingly different forms, the analysis appears to disregard this distinction by analyzing them together” (p. 122). Second, it raises the issue of accuracy being calculated based on “obligatory occasions” which does not take into account inaccurate usage. One argument in favor of keeping the same procedures in coding and data analysis is of course a higher degree of comparability. Yet, in line with McManus and Marsden's (2018) decision to include both sets of effect sizes in their replication, Porte and McManus do not exclude the possibility of “conducting both types of analysis followed by a comparison of findings” (2019, p. 123).

The discussion section of a replication should of course aim to “interpret contrasting sets of findings” (Porte & McManus, 2019, p. 141), and attempt to explain differences as well highlight



“commonalities” (p. 143) between the two studies. Similar findings between initial study and replication will often result in reinforcing our confidence in the initial hypotheses or expand the theory to additional context (for example in a replication that will use a different target language, similar findings between initial and replication studies could be an indication that the phenomenon being studied can be generalized to these two language contexts). Finally, researchers should address limitations (both in the replication study results, as well as the replication process) and make suggestions for future research, especially those that may call for additional variable modification (series of close or approximate replications).

In this first chapter, we discussed the state of replication in Applied Linguistics and the importance and value of replication in this field. We provided definitions for the different types of replications, and examined potential challenges and considerations researchers should pay attention to when undertaking replications. As mentioned in our introduction, our research project is based on a close replication of Taguchi et al. (2016). Some of the challenges we mentioned above are issues we faced during the different phases of our replication; therefore, it was important to mention them now. It should also be acknowledged that our study not only aims to replicate Taguchi et al. (2016) but also aspires to expand upon it. Consequently, some of our analyses and results may be seen as an extension of Taguchi et al. as well.

## Chapter 2 - Theoretical background

The goal of [Chapter 2](#) is to provide an overview of the concept of indirect meaning: what it means, how it is formed, and how people use it in general. In the first section we will provide a general definition of indirect meaning. The following sections will focus on different types of indirect meaning as well as various theories attempting to explain non-literal meaning production and comprehension processes.

### 2.1. Definition: what is indirect meaning?

Indirect meaning can take many forms: metaphors, idioms, hyperboles,<sup>3</sup> sarcasm, irony, as well as indirect speech acts (indirect refusals, indirect apologies, indirect requests), etc. Despite of all its forms and realizations in communication, a common—and all encompassing—definition is that indirect meaning presents a discrepancy between what is said and what is intended, the meaning that “goes beyond what is given by the language itself or what is literally said” (Verschueren, 1999, p. 25), what Bromberek-Dyzman and Ewert (2010) label the “communicative value.” Other terms authors have used to refer to the same idea include non-literal, figurative or non-natural meaning. Indirect meaning comprehension falls under the aspect of communicative competence called pragmatic competence, which, according to Taguchi (2007, p. 313-314) “refers to the ability to produce meaning in a socially appropriate manner and to interpret meaning, explicitly or implicitly stated, according to contexts.” Citing Grice (1975), Levinson (1983), Sperber and Wilson (1995) and Thomas (1995), Taguchi (2005) adds that pragmatic comprehension relies on “the ability to understand implied speaker intention by using linguistic knowledge, contextual clues, and the assumption of relevance” (p. 544). Thomas (1995) breaks down this pragmatic comprehension process into different levels: the abstract meaning of a sentence or utterance (“what a particular sentence could mean in theory” p. 16) goes through disambiguation (context is used to assign proper sense, reference and resolve possible sentence structure ambiguity) to reach the utterance meaning (“what the speaker actually does mean by these words on this particular occasion” p. 16). The third level of meaning is the force, or “the speaker’s communicative intention” (p. 18). While in some

---

<sup>3</sup> According to <https://literarydevices.net/adynaton/> adynaton is a rhetorical device that is a form of hyperbole in which exaggeration is taken to a great extreme where it seems impossible (e.g., “When pigs fly”). In other words, when hyperbole is magnified to such an extent that it is completely unfeasible, it is called adynaton.

cases the utterance meaning is a direct expression of the illocutionary force of the sentence (e.g., a direct request such as “Close the window”), sometimes, an illocutionary speech act (Austin, 1962) may take the form of an indirect meaning utterance (e.g., an indirect request: “It’s pretty cold in here, isn’t it?,” as a more subtle way to indicate that one would like the window to be closed because they feel cold). Speech acts that result in indirect meaning can range from apologies, indirect refusals, criticism, praise, request, etc. Thomas (1995) also specifies that the illocutionary force of an utterance may not always be obvious to the hearer (who may understand the speaker’s utterance but not the intended meaning). As stated by Taguchi (2005, p. 543):

Speakers make a variety of communicative choices, both verbal and non-verbal, which they enact for strategic reasons. The speaker’s choices may not always be transparent to the listener. Thus, it is an indispensable part of a listener’s pragmatic competence to recognize speaker intention conveyed in a variety of implicit and explicit forms and to make accurate inferences about a speaker’s communicative goals.

Sometimes however, the hearer may be able to understand the illocutionary force without understanding the utterance meaning, though Thomas (1995) points out, it is usually not as common. She provides an example in which the hearer (Alan) failed to understand the meaning of a word his girlfriend uttered and was thus unable to understand the utterance meaning. Yet, he was able to understand the force (his girlfriend was praising her dad) of the utterance. Thomas also refers to her personal experience hearing two idiomatic expressions (“It’s my shout” and “Don’t have a cow!”) that were said to her, and which she did not know (and was therefore not able to understand at the sentence/utterance level), but which she was nonetheless able to correctly interpret (using context) as an offer and a piece of advice (p. 20).

So how do people, native speakers or learners of a language, process this “mismatch between the literal utterance and the intention of the utterance [...] to comprehend the intention of the utterance” (Taguchi, 2005, p. 547)? To answer this question, we will first take a look at the theories of indirect meaning comprehension—since comprehension is the focus of our dissertation—which include Grice’s theory on Implicatures, as well as his Maxims of Conversation and Cooperative Principle (1975), and Sperber and Wilson’s Relevance Theory (1995). Special parts will also be dedicated to the theories related to politeness, as well as irony

production and comprehension. Finally, we will draw on the field of Conversation Analysis (CA) to take a closer look at how CA tackles indirect refusals and indirect opinions.

## 2.2. Implicature

According to Grice (1975, 1989) human communication relies on a search for the speaker's intended meaning. Therefore, when faced with indirect meaning, one must be able to move past the literal message of what is said to discover the actual meaning of the speaker's utterance, what he "implies." Grice's theory on *implicatures*—a term he coined to describe both the act of meaning, suggesting or implying ("implicating") one thing by saying something else, as well as the object of that act (Grice, 1975, pp. 43-44)—forged a path and serves as a paradigm for many theories of figurative meaning production and interpretation which we will discuss in this chapter.

The most commonly recognized forms of implicatures are figures of speech or tropes (metonymy, litotes, metaphor, etc.), but they may have a variety of communicative functions: indirect opinion, indirect refusal, indirect reply, etc. Grice makes a distinction between two types of implicature: *conventional implicature* and *conversational implicature*. Both convey an additional or different level of meaning "beyond the semantic meaning of the words uttered" (Thomas, 1995, p. 57). The main difference between the two however, is that conversational implicature is context-dependent and requires an inferential process that is unique to the situation of utterance (Bouton, 1994a, 1994b). Several examples of conversational implicatures will be provided and explained in the next section of this chapter, but for now let us consider the following example from Levinson (1983, p. 102):

- (1) A: Where's Bill?  
B: There is a yellow VW outside of Sue's house.

Taken literally, B's utterance in (1) does not answer A's question. Yet, using the context, the hearer may infer that Bill is at Sue's house (the implicature). As Taguchi states (2005, p. 545), conversational implicatures can be "understood only when context factors are taken into account." On the other hand, conventional implicature refer to formulations for which the implied meaning is based solely on the "conventional meaning of the words (the sentence) [the speaker] has uttered" (Grice, 1975, p. 44). Grice provides an example of conventional implicature in which the intended meaning was coded in the lexical item "therefore." When

saying “He is an Englishman; he is therefore brave” (p. 44), one implicates that being brave follows from being an Englishman. It is the meaning of the word “therefore” here that triggers the implicature. Other words that can activate conventional implicature include *but*, *even*, *yet*, and *for* (Levinson, 1983; Thomas, 1995). Davis (2019) argues that conventional implicatures can also be triggered by syntactic choices. For example, one could use appositive structures, which do not explicitly state the relationship between two words or expressions but rather implicate it. To illustrate this phenomenon Davis explains that by saying “Ravel, a Spaniard, wrote Spanish-style music” one implicates (but does not directly say) that Ravel was a Spaniard.

One key difference between conventional and conversational implicatures according to Grice (1975, 1989), is the fact that the latter should be:

- Cancellable:
  - explicitly: “by the addition of a clause that states or implies that the speaker has opted out” (Grice, 1975, p.57). In other words, denying the proposition or saying he did not mean it or adding some additional premises (Bromberek-Dyzman & Ewert, 2010).
  - contextually: “if the form of the utterance that usually carries it is used in a context that makes it clear that the speaker IS opting out” (Grice, 1975, p. 57).
- Calculable: language users should be able to calculate the speaker’s intended meaning based on what is said and the Conversation Maxims (which we will discuss in the next part).
- Non-detachable: the intended meaning is “attached to the semantic content of what is said and not to the linguistic form” (Bromberek-Dyzman & Ewert, 2010, p. 322). Implicatures are linked to a particular truth-conditional content (from which they cannot be detached).

Despite this distinction, several authors seem to disagree with Grice’s definition and use of “conventionality.” Indeed, relevance theorists (including Sperber and Wilson, which we will discuss shortly) consider that Grice’s conventional implicatures (as well as some generalized conversational implicatures) are actually explicatures and part of what is (pragmatically) said rather than implicatures (what is implied) (Carston, 1988; Gibbs, 2000, 2002). On the opposite end of the spectrum, Davis (1998, 2019) and Lepore and Stone (2015) argue that “even conversational implicatures can be conventional in the non-technical sense” (Davis 2019). Here “conventional” is used to refer to cultural norms, standard practices or social conventions

(though Davis makes a distinction between what he calls “semantic conventional implicatures,” as defined by Grice, and “customary conventional implicatures” to refer to the type of conventional implicatures defined above). This is consistent with Bromberek-Dyzman and Ewert’s (2010) definition of conventionality as being “related to the frequency of usage of a particular figurative expression and its salience in the speakers’ minds, which is relative to a community of usage” (p. 318). The concept of conventionality with respect to pragmatic competence, as we can see, seems to vary in definition. However, its role in comprehension occupies an important place in L2 pragmatic research. We will therefore come back to it during the literature review of L2 research ([Chapter 3](#)).

### 2.3. Cooperative Principle and Conversation Maxims

Wishing to explain and predict how implicatures can arise and be understood, Grice developed one of the most influential theories in the field of Pragmatics. According to Grice, the processes by which people derive the meaning of conversational implicature are ruled by the Cooperative Principle, which goes as follows: “make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose and direction of the talk exchange in which you are engaged” (1975, p. 45). Bromberek-Dyzman and Ewert (2010 p. 323) explain that the “inferences that arise to preserve the assumption of cooperation are implicatures.” Unless a hearer receives indications to the contrary, he will assume the speaker is being cooperative and when an utterance does not seem to fit with the logic of their conversation, the hearer will look for a different meaning in the speaker’s utterance. Thomas (1995) points out however, that this assumption of cooperative communicative behavior can sometimes prove to be wrong and that one may be deliberately misled by their interlocutor. However, she validates Grice’s attempt to generalize the pattern of communicative interaction as cooperative:

Grice was not (as some commentators have erroneously assumed) suggesting that people are always good and kind or cooperative in any everyday sense of that word. He was simply noting that, on the whole, people observe certain regularities in interaction and his aim was to explain one particular set of regularities — those governing the generation and interpretation of conversational implicature. (p. 63)

Along with the Cooperative Principle, Grice devised four conversational maxims (under which fall some supermaxims as well as specific maxims—1975, pp. 45-46) on which conversations are built:

Maxim of QUANTITY:

1. Make your contribution as informative as is required (for the current purposes of the exchange).
2. Do not make your contribution more informative than is required.

Maxim of QUALITY: “Try to make your contribution one that is true”

1. Do not say what you believe to be false.
2. Do not say that for which you lack adequate evidence.

Maxim of RELATION: “Be relevant”

Maxim of MANNER: “Be perspicuous”

1. Avoid obscurity of expression.
2. Avoid ambiguity.
3. Be brief (avoid unnecessary prolixity).
4. Be orderly.

In conversation, one has the option to observe all of these maxims. Thomas (1995, p. 64) provides an example of such interaction:

- (2) Husband: Where are the keys?  
Wife: They're on the table in the hall.

What the wife said as a response to her husband's question is a direct and explicit message providing the location of the keys. Her answer is not an implicature (there is no difference between what she said and what she meant) and she followed all four maxims: she spoke clearly (manner) and truthfully (quality) and gave enough—but not too much—information (quantity). Her answer was in direct relation to her husband's question (relation).

However, Grice explains that a speaker can purposely decide not to abide by a maxim, and in the case of conversational implicatures specifically, Grice states that (1975, p. 49):

He [the speaker] may FLOUT a maxim, that is, he may BLATANTLY fail to fulfill it. On the assumption that the speaker is able to fulfill the maxim and to do without violating another maxim (because of a clash), is not opting out, and is not, in view of the blatancy of his performance, trying to mislead, the hearer is faced with a minor problem: How can his saying what he did say be reconciled with the supposition that he is observing the overall CP? This situation is one that characteristically gives rise to a conversational implicature; and when a conversational implicature is generated in this way, I shall say that a maxim is being EXPLOITED.

Grice also provides examples of conversational implicatures that may deliberately choose not to observe specific maxims. When using verbal irony for instance, the speaker, according to Grice, can exploit the maxim of Quality. Grice (1975) illustrates the phenomenon with the following situation: after being betrayed by X who revealed his secret to a competitor, A says ‘X is a fine friend’ (p. 53). In this situation, the audience is aware that A has said something that he does not believe is true, and the audience also knows that A is aware that this is obvious to his audience. By assuming that A follows the CP, this apparent flouting of the maxim of Quality leads the hearer to seek a different interpretation by inferencing that what A “meant” was actually the opposite of what he said.

Let us now consider the following example (excerpt from our study):

- (3) Jacques: Hi, Sandrine! It is a beautiful day to work out outside today.  
Sandrine: Yes, it’s the perfect weather to go jogging in the park.  
Jacques: Do you want to go jogging with me this afternoon?  
Sandrine: I already told my friend I would play tennis with her.

In (3) Sandrine does not directly answer Jacques’ question. What she says (her explicature) is that she previously told one of her friends that she would be playing tennis with her. But the hearer, Jacques, is meant to infer that she cannot go jogging with him that afternoon. The implicature here is an indirect refusal of Jacques’ invitation, and it arises from the maxim of Relation. If Sandrine was being cooperative, then we can only assume that she was trying to answer Jacques’ question and the hearer can work out the incompatibility between his request (jogging) and the plans Sandrine already made, and infer that she is unable to go jogging with him. As Taguchi et al. state, “indirect meaning is understood based on the assumption that the speaker operates under the cooperative principle, as well as the listener’s ability to supply contextual information to make inferences from seemingly unrelated utterances” (2016, p. 679).

Grice (1989) and Paltridge (2006) differentiate between maxim *violation* and maxim *flouting*: with the former, the speaker manipulates the meaning of the sentence to mislead the interlocutor. For example, a mother might tell her child that “Mummy’s gone on a little holiday because she needs a rest” (Paltridge, 2006, p. 65) instead of admitting that she is separating from her husband. She *violates* the maxim of Quality in the sense that she is not being completely truthful—she is lying—about a potential divorce. On the other hand, Paltridge explains that when a speaker is *flouting* a maxim, he is not trying to mislead the hearer but



rather, he is deliberately not following the maxim to lead his hearer to a different intended meaning (see example of indirect refusal above (2)).

In addition to maxim violation and flouting, Grice also considers the possibility that a speaker may “*infringe*” on a maxim. However, this is usually done unintentionally: either because the speaker’s command of the language is imperfect (child or L2 learner) or because of some kind of impairment (cognitive, speaking disability, performance issue—nervousness, excitement, etc.—inebriety, and so on).

Sometimes speakers are simply not able or allowed to provide information or answer questions. They can then “*opt out*” of the conversational maxims. This does not necessarily mean that they are being uncooperative, but Thomas (1995) explains that speakers may not be able to share information for confidentiality reasons (a priest, therapist, doctor, etc.) or because it might hurt or put someone in danger.

Some authors also argue that Grice is missing a fifth situation in which a speaker may not abide by the Conversational Maxims. In this case, they talk about “*suspending a maxim*” (Thomas, 1995, p. 76). Thomas (1995) and Bouton (1988, 1994a, 1994b) reference the work of Keenan (1976) whose research with the Malagasy people highlights the importance of culture in the application of the conversational maxims. Indeed, Keenan (1976), trying to prove the limitation of Grice’s conversational implicature’s theory, states that Malagasians “regularly provide less information than is required by their conversational partner, even though they have access to the necessary information” (p. 70). However, Thomas (1995) argues that in Malagasy society, there is no expectations that speakers will provide all relevant information they possessed. As such, they are not being uncooperative, but just *suspend* the maxim of Quantity as it is the norm among members of their community. While imperfect, Grice’s theory does account for the fact that the speaker and hearer should enter a conversation on some common grounds that will make the implicature comprehension possible, such as “the **roles and expectations** of the participants in a conversation [...] the **context** in which the utterance occurs; and [...] the **world around them** as it pertains to their interaction” (as paraphrased in Bouton, 1994b, p. 89). Summarizing Grice’s ideas regarding the process by which the hearer will work out a particular conversational implicature, Davis (2019) writes:

[...] the hearer will rely on the following data: (1) the conventional meaning of the words used, together with the identity of any references that may be involved; (2) the Cooperative Principle and its maxims; (3)

the context, linguistic or otherwise, of the utterance; (4) other items of background knowledge; and (5) the fact (or supposed fact) that all relevant items falling under the previous headings are available to both participants and both participants know or assume this to be the case.

For communication to be successful, all conditions (assume that the speaker is cooperative, believes in the intended message, and mutually shared knowledge and perspective) must be met.

While Grice's theories of implicature comprehension still serve as theoretical framework for both L1 and L2 research, they were not without flaws. First, Grice only considered implicature as part of an interaction in which two or more interlocutors communicate and negotiate meaning. However, Davis (1998) points out that implicature (in the form of various figures of speech) could also be used in private journals or manuscripts. These would still exist despite the absence of a reader, which means that a communicative interaction does not have to take place for an implicature to be produced (even if no one would then interpret it). As previously mentioned, some authors also argue that Grice's overall assumption that the speaker wants to be cooperative by default is limitative. Indeed, speakers may engage in verbal communication without abiding by the Cooperative Principle, and following the Cooperative Principle is not a requirement for the existence of an implicature. Drawing from the field of evolutionary social psychology, and more specifically game theory, Pinker (2007) explains that expecting perfect and constant cooperation in communication is unrealistic. Furthermore, Grice's Cooperative Principle does not consider the fact that someone may lie or purposely withhold information they know to be true. Davis (2019) provides the following example, showing that a speaker may still use a conversational implicature while violating the Cooperative Principle. He says "[i]f the prosecuting attorney asks the defendant whether he was in the bank on the day of the robbery, the defendant might answer "I took my mother to the hospital that day," thereby implicating that he was not at the bank. He may do this in the hope of misleading the jury as well as the prosecutor." The defendant is not being truthful (not cooperative in his goal), but the interpretation of his utterance still may follow the mechanisms of a conversational implicatures (cf. the difference between maxim violation and maxim flouting from Paltridge, 2006). Another example of the limitations of the Cooperative Principle is jokes (Lepore & Stone, 2015): when a speaker makes a joke, they are not trying to be informative: their main goal may be amusement. In that instance, the maxims of Quality, Quantity and Relation cannot be applied. Even Grice recognizes his maxims may "clash" (1975, p. 30). As a result, some

authors have tried to expand on Grice's Cooperative Principle, and add additional Principles that may explain conversational implicatures with goals other than communication efficiency (the latter being central to the Cooperative Principle). We will now briefly present a few of these complimentary Principles that built upon Grice's theory.

Davis (2019) introduces the Principle of Style ("Be stylish, so be beautiful, distinctive, entertaining, and interesting," also see Tannen, 1987) to account for the ways people may express themselves that would conflict with the maxim of Manner (the speaker may use some figures of speech and loose perspicuity to liven up their story), Quality (the speaker may "embellish their story to make it more interesting") and/or Quantity (the speaker may omit "boring or ugly details").

Another notable effort to resolve the issues of Grice's principle clashes came from Horn (1984) who proposed two interrelated principles to replace Grice's maxims of Quantity, Relation and Manner, while retaining Quality as having a "privileged status" without which other maxims cannot be satisfied (Horn, 2004, p. 7).

- **Q-principle:** Make your contribution sufficient; say as much as you can (given R).
- **R-Principle:** Make your contribution necessary: say no more than you must (given Q).

The Q-Principle parallels Grice's first maxim of Quantity ("Make your contribution as informative as required") as well as the first two maxims of Manner (avoid obscurity and ambiguity), while the R-Principle combines Grice's Principle of Relation ("Be relevant"), the last two maxims of Manner (be brief and orderly), and the second maxim of Quantity ("Do not make your contribution more informative than is required").

Consider the following examples, frequently cited as giving rise to Q-based implicatures (4) and R-based implicatures (5):

- (4) a. He broke into a house.
- b. He did not break into his own house.
- (5) a. I broke a finger yesterday.
- b. I broke one of my fingers yesterday.

In Horn's view, since the speaker of (4a) does not provide specific information about the house, the hearer should infer (4b) as deriving from the Q-Principle. On the other hand, the use of the indefinite article in (5a) should be interpreted on the basis that the speaker provided the minimum amount of information necessary to get their point across. Horn's overarching idea centers around the conflicting interests of the speaker and those of the hearer: The Q-Principle thus represents the pressure that the hearer may place on the speaker, while the R-Principle

favors the minimization of production efforts from the speaker. According to Horn (1984, 2004), the fact that these two principles may conflict are at the basis of possible miscommunication issues since the speaker and hearer may be exploiting or focusing on different principles. However, Horn's critics point out that while Horn did identify two patterns of interpretation, he fails to explain how these particular implicatures are reached. Levinson (2000 but also in Atlas & Levinson, 1981), who adopted similar principles which he labelled Q and I (instead of R—I standing for Informativeness) reintroduced the principle of Manner (Levinson, 2000, p. 136-137):

- **M-Principle:** Indicate abnormal, non-stereotypical situation by using marked expressions that contrast with those you would use to describe the corresponding normal, stereotypical situation.

He also indicates that in order to avoid the clashes from Grice's maxims, there should be an order of preference by which the Q-Principle is applied first, followed by the M-Principle and then the I-Principle. Nevertheless, other researchers such as Huang (2014) and Davis (2019) disagree with the proposed ordering.

The above is just a short overview of some of the neo-Gricean pragmatics theories of implicatures that followed Grice's Cooperative Principle. We have opted to devote a more substantial section to the theories related to politeness, which also stemmed from Grice's Cooperative Principle and related Maxims. The face-saving or politeness goal of some conversational implicatures is particularly important for this dissertation because the three types of indirect meaning we will focus on—indirect refusals, indirect opinions and irony—can be tied to it, either because they may subscribe to or build upon it, or on the contrary, exploit or go against it.

#### 2.4. Theories of Politeness (and related principles)

Leech (1983, 2005, 2014) and Brown and Levinson (1978, 1987) (and before them Lakoff, 1973) focus on politeness as a pragmatic aspect of communication where politeness is "interpreted as a strategy (or series of strategies) employed by a speaker to achieve a variety of goals, such as promoting or maintaining harmonious relations" (Thomas, 1995, p. 158). As such, politeness plays a crucial part in interaction "because it controls the cooperative behavior in conversation" (Al-Duleimi, Rashid, Sabariah and Abdullah, 2016).

In his 1983 book, *Principle of Pragmatics*, Leech introduces the ***Principle of Politeness*** which, he postulates, can be seen “not just as another principle to be added to the CP, but as a necessary complement, which rescues the CP” (1983, p. 80) by providing a reason as to why people might be—on the surface—uncooperative. The Principle of Politeness is presented as having two forms: a negative version (“Minimize (other things being equal) the expression of impolite” p. 81) and a positive one (“Maximize (other things being equal) the expression of polite beliefs” p. 81). To accompany the Principle of Politeness (and in much of the same way Grice’s presented his), Leech also introduces six interpersonal maxims (1983, p. 132):

*Tact*: Minimize cost to other; maximize benefit to other.

*Generosity*: Minimize benefit to self; maximize cost to self.

*Approbation*: Minimize dispraise to other; maximize praise of other.

*Modesty*: Minimize praise of self; maximize dispraise of self.

*Agreement*: Minimize disagreement between self and other; maximize agreement between self and other.

*Sympathy*: Minimize antipathy between self and other; maximize sympathy between self and other.

These maxims could explain why, when asked for one’s opinion, one might decide to answer indirectly so as to minimize cost or dispraise to other. The Politeness Maxims will in essence override the Conversational Maxims (Bouton, 1884). Leech provides the following two examples (among many other) to illustrate his Principle of Politeness (1983, p. 80)

(6) A: We’ll miss Bill and Agatha, won’t we?

B: Well, we’ll all miss BILL.

(7) P: Someone’s eaten the icing off the cake.

C: It wasn’t ME.

In (6) the Maxim of Quantity is violated since B does not seem to answer the question completely and ignores part of the question. The implicature here is that B’s opinion is that they will not miss Agatha. Of course, B could have explicitly stated this (“We will miss Bill but not Agatha”) following the Cooperative Principle. However, this addition (which would have provided all the requested information) would have likely resulted in B being perceived as quite impolite. B’s indirect message about not missing Agatha could still be reached but on the surface, he remained tactful by not saying something negative about Agatha. He “suppressed the desired information in order to uphold the PP [Politeness Principle]” (Leech, 1983, p. 80).

Example (7) looks like a typical conversation between a parent P and a child C. C's apparent irrelevant reply to P's remark stems from the fact that P's remark is an implicature and that without directly accusing C, P may have implied that C could be the one who ate the icing of the cake. Leech's explanation for this apparent breach of the Maxim of Relation is as follows (1983, p. 81):

Suppose *P* is not sure who is the culprit, but suspects that it is *C*. Then a small step of politeness of *P*'s part would be to withhold a direct accusation, and instead to make a less informative, but undoubtedly true assertion, substituting an impersonal pronoun someone for the second-person pronoun you. Thus, *P*'s remark in [6] is interpreted as an indirect accusation: when *C* hears this assertion, *C* responds to it as having implicated that *C* may well be guilty, denying an offence which has not been overtly imputed. [...] *C* responds to that implicature, the indirectness of which is motivated by politeness, rather than to what is actually said.

Leech also notes that the two examples above (5 and 6), can “tip over into an ironic interpretation” (p. 82). He therefore proposes a “second-order principle which builds upon, or exploits, the principle of politeness: The *Irony Principle*, which takes the following form (p. 82): “If you must cause offence, at least do so in a way which doesn't overtly conflict with the PP, but allows the hearer to arrive at the offensive point of your remark indirectly, by way of implicature.” The Irony Principle “enables a speaker to be impolite while seeming to be polite; it does so by superficially breaking the CP but ultimately upholding it” (p. 142). This “insincere politeness” can appear when the speaker flouts the maxim of Quantity or Quality. At the opposite end of this spectrum, Leech's “*Banter Principle*” gives rise to implicatures that show solidarity or intimacy to the hearer by pretending to be offensive: “what the speaker says is impolite to the hearer and is clearly untrue. Therefore, what the speaker really means is polite to the hearer and true” (1983, p. 144). We can see here that Leech sees both irony and banter from the perspective of meaning reversal (Jobert & Sorlin, 2018) where one mirrors the other: “irony maintains courtesy on the surface level of what is said, but at a deeper level is calculated to imply a negative evaluation. Banter is offensive on the surface but at a deeper level is intended to maintain comity” (Leech, 2014, p. 100). We will see in our section on Irony (2.6) that, while many authors still rely on this definition of meaning reversal, irony can take on many other forms.

While Leech's theories were anchored in sets of principles and maxims participants in a communicative exchange were intended to follow (or as Thomas puts it “a series of social-

psychological constraints influencing, to a greater or lesser degree, the choices made within the pragmatic parameters” 1995, p. 168), another influential theory of politeness that focused more on the concept of “face” was put forward by Brown and Levinson in 1978 and later revised in 1987 (Thomas, 1995). They based their theory on Goffman’s (1955) ideas which define the term “face” as:

[...] the positive social value a person effectively claims for himself by the line others assume he has taken during a particular contact. Face is an image of self-delineated in terms of approved social attributes — albeit an image that others may share, as when a person makes a good showing for his profession or religion by making a good showing for himself. (Goffman, 1967, as quoted in Thomas, 1995, p. 168)

Brown and Levinson postulate that some speech acts can present a threat to other’s or one’s face (essentially damaging the image of the hearer, the speaker, or someone outside of the conversation). Three factors come into play and should be consider during an interaction: *power* (social status), *social distance* (degree of familiarity between the speaker and hearer) and *degree of imposition* (rank of the speaker and ability to impose their ideas and desires onto others). All three parameters are context- or culture-dependent according to Brown and Levinson. Based on these, a speaker can evaluate the risk that might be cause to their interlocutor’s face and choose their utterance accordingly. Brown and Levinson further define the concept of face as follow (1978, p. 62):

**negative face:** the want of every ‘competent adult member’ that his actions be unimpeded by others.

**positive face:** the want of every member that his wants be desirable to at least some others.

A speaker may opt to perform a face-threatening act (at the risk of sounding bold or rude), or employ strategies that will preserve their interlocutor’s face. Potentially face-threatening acts to the hearer’s negative face include (Brown & Levinson, 1978, p. 66): utterances that put pressure on the hearer to do or refrain from doing something (e.g., orders, requests, suggestions, threats, warnings, reminders, etc.), utterances that put pressure on the hearer to accept or reject a future action from the speaker, which could put the hearer in his debt (e.g., offers and promises), utterances that may lead the hearer to want to give something to the speaker or on the contrary protect him against it (e.g., compliments, expressions of envy, anger, lust, etc.). Face-threatening acts that may challenge the hearer’s positive face function (p. 66) “by indicating (potentially) that the speaker does not care about the addressee’s feelings, wants, etc. — that in some important respect he doesn’t want H’s wants” and can be seen in some of the

following:

- utterances that show that the speaker has a negative opinion of the hearer (e.g., expressions of disapproval, criticism, ridicule, complaints, reprimands, accusation, etc.)
- utterances that show the speaker's indifference to the hearer's feelings (e.g., expressions of violent emotions, irreverence, willingly causing distress to the hearer by bringing up bad news, divisive topics, etc.).

Brown and Levinson explain that on some occasions, speakers may not have a choice but to speak directly. They called performing a face-threatening act without any redress "*bald on record*" (1978, p. 94). This strategy conforms to Grice's Maxims as the speaker is direct, clear, truthful and provides the appropriate amount of information needed for the conversation. It may be that the speaker is facing an emergency and must get their message across quickly (Thomas, 1995). In that instance, efficiency of communication prevails over politeness and "no face redress is necessary" (Brown & Levinson, 1978, p. 95).

Nonetheless, a speaker may also use some face-saving strategies to reduce the threats on the hearer's positive or negative face (in that case, they are said to use "*redress*"). Brown and Levinson speak of *positive politeness* to describe speech acts that are directed to the hearer's positive face (i.e., the speaker takes the hearer's wants, desires and rights into account and tries to satisfy them) and *negative politeness* (which focuses on the hearer's want to be free to act as they desire). Brown and Levinson provide a list of strategies that can help save the hearer's positive face with redress (positive politeness): attend to the hearer's interests, wants, needs; sympathize, show approval, use in-group identity markers or in-group language/dialect, seek agreement, avoid disagreement, hedging opinions (i.e., be vague/e.g., using specific discourse marker like the expression "sort of" to soften the face-threatening act), offers, give gifts, and jokes. The latter, Brown and Levinson state, are "based on mutual shared background knowledge and values, jokes may be used to stress that shared background or those shared values. Joking is a basic positive-politeness technique, for putting H 'at ease'" (1978, p. 124). Here it is clear that they are considering jokes not for their potential ability to make fun of someone else but as way to create a sense of complicity between hearer and speaker or to make difficult situations seem less tense (we will come back to this aspect in our section on irony below).

Negative politeness is usually linked more strongly to the social distance factor we discussed earlier and as such can be expressed through more conventionalized expressions because it



includes the use of linguistic forms that show “respect, deference, and politeness” (Kida, 2011, p. 183). Strategies that allow a speaker to reduce the threats to the hearer’s negative face with redress include (Brown & Levinson, 1978): use of “conventional indirectness” (i.e., routine idioms or formulas/e.g., “Can you please pass the salt?” instead of using a direct expression with an imperative) which can vary in degrees of politeness,<sup>4</sup> hedging or questioning, minimizing the imposition, deference (i.e., the speaker may adopt a humble stance or abase themselves or on the contrary, the speaker might raise the hearer’s status), apologizing, impersonalizing (e.g., use of impersonal structures or passive voice).

In addition to threat to negative and positive face, a third type of face threatening act has been identified by Brown and Levinson: “off record politeness” is done in a way that leaves room for several interpretations. Their definition reminds us of Grice’s conversational implicature definition as it allows the speaker an “out” by “providing himself with a number of defensible interpretations” (Brown & Levinson, 1978, p. 211). Indeed, the speaker may imply something but not say it directly which results in the implicature being “cancellable.” Brown and Levinson continue by explaining that “to construct an off-record utterance one says something that is either more general (contains less information in the sense that it rules out fewer possible states of affairs) or actually different from what one means (intends to be understood)” (p. 211). It is very clear from this statement that the utterance would then violate Grice’s Maxims, which Brown and Levinson consider as being the trigger for the hearer to realize that an inference must be made (p. 211 and p. 213<sup>5</sup>). Without going into too many details, Brown and Levinson mention that the type of inference used by the hearer to recover the intended meaning is a “matter of contention” (p. 211) but they cite various authors who favor specific types of strategies: Lakoff (1990) and semantic entailment, Searle (1975) and inductive reasoning, Atlas and Levinson (1981) and practical reasoning. A detailed account of the various sub-strategies one might use to produce off-record politeness is presented in Brown and Levinson (1978), in the form of a chart showing which of Grice’s Maxims each strategy may violate (p. 214). On

---

<sup>4</sup> “Can you pass the salt?” is considered polite but most maybe less so than “Could you pass the salt?.” “Could you possibly pass the salt (please)?” may seem like the politest expression of the same speech acts. However, authors (Brown & Levinson, 1978; Leech, 2014; Thomas, 1995) explain that the degree of politeness perceived is also culture and context-dependent. For example, the third sentence uttered by a wife to her husband may actually convey feelings of frustration and anger, because this level of formality and politeness is not usually called for between two people who know each other.

<sup>5</sup> Brown and Levinson, 1978, p. 213: “If a speaker wants to do an FTA, and chooses to do it indirectly, he must give H some hints and hope that H picks up on them and thereby interprets what S really means (intends) to say. The basic way to do this is to invite conversational implicatures by violating, in some way, the Gricean Maxims of efficient communication.”

the one hand, the speaker may opt to be vague and ambiguous (violating the Manner Maxim). On the other hand, the speaker may rely on conversational implicatures triggered by the violation of the following Maxims:

- Relevance/Relation Maxim (hints, association clues, and presupposition)
- Quantity Maxim (understatement, overstatement, and tautologies)
- Quality Maxim (contradictions, irony, metaphors, and rhetorical questions).

Brown and Levinson note that the implicature strategies are often accompanied by clues leading to their interpretation (“the mutual knowledge of S and H in the context; the intonational, prosodic and kinesic clues to speaker’s attitude; the clues derived from conversational sequencing, add up to only one really viable interpretation in the context” p. 212).

We have reviewed the main theories of implicature comprehension that include Grice’s Cooperative Principle and Conversation Maxims, as well as a few other attempts to work out some of their limitations. We have delved into two of the most influential theories of Politeness and derived Principles. By now, it should be obvious that none of the above theories were perfect: as a matter of fact, some of the cited authors even looked into improving upon their own theories (as it is the case with Leech who, in 2005, present his “Grand Strategy of Politeness,” in response to criticisms from authors saying that his Theory of Politeness was not suitable to Eastern Cultures and Languages). In the next section ([2.5](#)), we will discuss Sperber and Wilson’s Relevance Theory, which is central to many L1 and L2 comprehension research studies (including the study we are replicating, Taguchi et al., 2016).

## 2.5. Relevance theory

Unsatisfied with Grice’s conversational maxims, Sperber and Wilson (1986, 1995) argue that “Grice’s maxims can be replaced by a single principle of relevance—that the speaker tries to be as relevant as possible in the circumstances—which, when suitably elaborated, can handle the full range of data that Grice’s maxims were designed to explain” (Wilson & Sperber, 1986, p. 381). This inferential processing account of human communication is based on the following assumptions: (Wilson & Sperber, 1998, p. 8):

First, every utterance has a variety of possible interpretations, all compatible with the information that is linguistically encoded. Second, not all these interpretations occur to the hearer simultaneously: for

example, some disambiguations, some contextual assumptions, some implicatures require more effort to recover. Third, hearers are equipped with a single, very general criterion for evaluating interpretations as they occur to them. And, fourth, this criterion is powerful enough to enable the hearer to recognize the intended interpretation as soon as it is encountered, without having to construct and evaluate a range of alternative interpretations.

The criterion proposed in *Relevance* is based on a fundamental assumption about human cognition: that human cognition is relevance-oriented; we pay attention to information that seems relevant to us. Now every utterance starts out as a request for the hearer's attention. As a result, it creates an expectation of relevance. It is around this expectation of relevance that our criterion for evaluating possible interpretations is built.

Sperber & Wilson define *Relevance* in terms of cognitive effects and processing effort: the greater the cognitive effects (when new information interacts with existing assumptions by either strengthening them, contradicting or eliminating them—cf. Wilson & Sperber, 1998—or leading to a new contextual implication—i.e., contextual effects—cf. Wilson & Sperber, 2007), the greater the relevance will be. However, the mental effort that comes from processing this information can be so great that it will lower the relevance of the utterance (and can put the speaker at risk of losing the hearer's attention). From this follows a *Cognitive Principle*—or *Principle of Maximal Relevance* (i.e., Cognitive Efficiency):

[H]umans automatically aim at maximal relevance: that is, maximal contextual effect for minimal effort. This is the single general factor which determines the course of human information processing. It determines which information is attended to, which background assumptions are retrieved from memory and used as context, which inferences are drawn.

(Wilson & Sperber, 2007, p. 47).

The second principle of relevance is the *Communicative Principle*, which leads to a presumption of *Optimal Relevance* that accounts for the speaker's abilities and preferences:

[T]he communicator is manifestly limited by her own abilities (to provide appropriate information, and to present it in the most efficient way). Nor can she be expected to go against her own preferences (e.g., against the goal she wants to achieve in communicating, or the rules of etiquette she wishes to follow).

(Wilson & Sperber, 2012, p. 64-65)

Considering the speaker's preferences through the Communicative Principle allows Sperber and Wilson to encompass previously labelled Principles of Politeness and Style. Processing of

indirect meaning therefore relies on the ability to not only decode the linguistic complexity of an utterance, but also to attend to contextual cues (such as “encyclopaedic entries [which] may contain ready-made chunks or schemas describing often-encountered [situations]”—Wilson & Sperber, 1998, p. 14), as well as accounting for the speaker’s intention and preferences. Sperber and Wilson break down the comprehension process into three sub-tasks (Wilson & Sperber, 2002, p. 261):

- a) Constructing an appropriate hypothesis about explicit content (in relevance-theoretic terms, EXPLICATURES) via decoding, disambiguation, reference resolution, and other pragmatic enrichment processes.
- b) Constructing an appropriate hypothesis about the intended contextual assumptions (in relevance-theoretic terms, IMPLICATED PREMISES).
- c) Constructing an appropriate hypothesis about the intended contextual implications (in relevance-theoretic terms, IMPLICATED CONCLUSIONS).

They also note that the comprehension process is not a sequential one and all three of the above hypotheses can be developed “in parallel against a background of expectations (or anticipatory hypotheses) which may be revised or elaborated as the utterance unfolds” (p. 262). This is important because it departs from the two or multi-stage processing model that Grice and the neo-Gricean theorists considered for the comprehension of indirect meaning (i.e., the obligatory processing of the literal meaning before accessing the intended, indirect meaning). We may remember here what Thomas (1995) had said about being able to understand the illocutionary force of an utterance without necessarily understanding it at the sentence-level (e.g., “It’s my shout” and “Don’t have a cow!”). This is further evidence that a hearer may still attempt to process the literary meaning of an utterance but even if he does not reach understanding, he might still be able to comprehend the intention of the speaker and what he implicated. Thomas states (1995, p. 21) that: “It is frequently (but not invariably) the case that we derive force from utterance meaning, but we can, for example, use paralinguistic features (such as intonation and tone of voice) or non-linguistic features (such as gesture) in order to work out the intended force.”

The issue of the literal processing of indirect meaning utterance is central to research that has been focusing on irony comprehension. Much has been written about the production and interpretation of verbal irony, sarcasm and humor, and it seems that to this day, authors do not always agree on how this pragmatic phenomenon is processed. Our next section will focus on

what makes irony so special and how different authors (from Grice, to Sperber and Wilson, as well as researchers that prescribe to more socio-cognitive and psycholinguistic approaches) understand the mechanisms of its comprehension. It should also be noted that we believe that what has been said about irony processing and comprehension, in many cases, can be (and has been) expanded and applied to other forms of indirect meaning, like indirect opinions (which are sometimes expressed in the form of ironic utterance) and indirect refusals.

## 2.6. Indirect meaning processing: The case of irony

Our dissertation is investigating L2 comprehension of three different types of indirect meaning: indirect refusals, indirect opinions and irony. While the theories (Grice's Cooperative Principle, the Relevance Theory and Theories of Politeness) we presented above shed some light on the production and processing of different kinds of implicatures, including indirect refusals and indirect opinions (and Taguchi has studied these two forms of indirect meaning in a series of replication and extension studies—cf. Taguchi 2002, 2005, 2007, 2008a, 2008b, 2008c, 2009a, 2009b, 2011; Taguchi, Li, & Liu, 2013), irony is a pragmatic device that has given rise to a large number of studies by L1 researchers, and is still the source of debate and new theories. For this reason, we wanted to dedicate a specific section to irony and how it is processed.

While most people usually have an idea of what irony is, the concept is actually hard to define. As Jobert and Sorlin (2018, p.3) state:

Call it ironic or not, the study of irony has inspired an impressive number of articles and books. Yet theoreticians are still fighting over where to place its defining borders. With each new paper or book come novel theoretical nuances accounting for various aspects of what turns out to be a multi-faceted phenomenon.

Attardo (2002) explains that the definition of humor, irony and sarcasm is problematic in that there is quite a bit overlap and the difference between these terms is not always clear. He defines irony as a “purely pragmatic phenomenon, without semantic counterpart. What this means is that the semantics of an ironical sentence and of a non-ironical sentence are indistinguishable. Naturally, in context, an ironical sentence will acquire (inferentially) a meaning that is minimally different and in fact in most cases opposed to that which the sentence would have in a ‘neutral context’.” (p. 163). According to him, there are many processing theories when it comes to irony production and comprehension (such as direct access, two-stage theory, graded

salience, etc., which we will discuss shortly), but the key to irony is the contrast and incongruity between an actual situation and the expectations and/or utterance, which he refers to as “inappropriateness.” Along the same line, Gibbs, Bryant, and Colston (2014) state that “irony is an umbrella term covering a variety of more specific figurative intentions, including sarcasm, jocularity, rhetorical questions, hyperbole and understatement (Gibbs 2000)” (pp. 576-577).

In this section, we propose to examine the different forms of irony and how they are produced in order to gain a better understanding of how people can detect and understand them. We will review the different theories of irony processing—including the role of literal meaning comprehension—as well as previous research focusing on the functions and effects of verbal irony.

### 2.6.1. *Definition of irony*

Verbal irony is a complex pragmatic (and paralinguistic) phenomenon that has been the subject of many theories. It is often hard to distinguish irony from sarcasm or even humor, and to decide if one is just a form of the other category. The traditional definition of irony that one can find in many dictionaries (French or English language) often refers to a direct opposition between what is said and what is actually meant: “Irony is a subtle form of humor that involves saying things that are the opposite of what you really mean.” (<https://www.collinsdictionary.com/us/> - in French, we can also cite <https://www.le-dictionnaire.com/> “*une figure de rhétorique par laquelle on dit le contraire de ce qu'on veut faire entendre*”). However, as Eggs (2009) highlights, the classic definitions from Du Marsais (1730)<sup>6</sup> and Fontanier (1830) where the intended meaning (“*le signifié*”) is the contrary of what is said (“*le dit*”), fail to account for all the instances of verbal irony. Consider the following example (8), cited by several authors (Colston, 2000, 2002, 2017; Colston & Gibbs, 2007; Colston & O’Brien, 2000; Gibbs, 1986, 2000; Gibbs et al., 2014, Sperber & Wilson, 1995) that illustrates how an ironic utterance can keep its truth value all the while signifying something different, but not necessarily the contrary or polar opposite of the literal utterance.

- (8) A taxi drivers says «I sure love drivers who signal before passing! », when another driver just cut him off without signaling before changing lanes.

---

<sup>6</sup> Dumarsais, 1730, p.156 : « une figure par laquelle on veut faire entendre le contraire de ce qu'on dit : ainsi les mots dont on se sert dans l'ironie, ne sont pas pris dans le sens propre et littéral »

Here the taxi driver simply wants to express his frustration with the bad driver. What is said is not the contrary of what is meant since the speaker is not actually implying the opposite of the literal utterance (which would have been “I hate drivers who signal before passing”). The literal interpretation (the fact that the taxi driver appreciates drivers who abide by the rules and signal) is not questioned. However, it is the contrast between what is said and the situation or context that allows the hearer to interpret the utterance (“*le dit*”) as an ironic implicature. Several authors will therefore use the word “contrast” rather than “contrary” to define irony. This contrast, or incongruity when referring to irony, humor and/or sarcasm (Attardo, 2002; Attardo & Raskin, 2017; D’imperio, Champagne-Lavau, Leovenbruck, & Fouillaud, 2013; Haugh, 2017; Ivanko & Pexman, 2003; Larking-Galiñanes, 2017; Shively, 2013; Yus, 2000) or even illogism (« inappropriateness » in Attardo, 2000a, 2002, 2017a, 2017b) can be found between the utterance (“*le dit*”) and its interpretation (“*le signifié*”), but also between the utterance and its context/cotext (Attardo, 2000b; Bres, 2010 ; Burgers, van Mulken, & Schellens, 2011). Therefore, definitions of irony, sarcasm or even humor often overlap. Colston (2017) for example, considers that sarcasm, ironic praise and criticism, hyperbole, ironic analogy and reformulation are all forms of verbal irony. Charaudeau (2006, 2011) characterizes one of his humor categories (“*l’humour par jeu énonciatif*”) as verbal irony as well (“*ironie énonciative*”). Jorgensen (1996) also speaks of “sarcastic irony,” thus blurring the lines between sarcasm and irony. A study by Bossut (2011, based on Channon, Rule, Maudgil, Martinos, Pellijeff, Frankl, Drury, & Shieff, 2007) surveying French native speakers, revealed that French people often consider the meaning and functions of sarcasm to be very close to what we would also define as irony. She lists the following as definitions provided by her participants (2011, p. 77—my translation):

- say the opposite of what one thinks/ ironic
- making fun, criticizing, negative expression
- humor
- soften/lessen impact of utterance (more polite)
- bonding between interlocutors (complicity)

Another element of the use of irony put forth in the traditional definitions, is the fact that irony is usually used to express negative feelings (Larousse.fr : « *manière de railler, de se moquer* »). Yet again, many authors, like Berrendonner (1981) and Charaudeau (2006, 2011) show that asteism (saying something negative to express a positive feeling), is also a form of verbal irony, even though, as others have highlighted (Clark & Gerrig, 1984; Colston, 2000; Gibbs, 2000; Kumon-Nakamura, Glucksberg, & Brown, 1995; Sperber & Wilson, 1981; etc.), there is

a clear asymmetry in its use as people will more often rely on irony to express negative feelings rather than positive. We already know that Leech's Principle of Irony and Principle of Banter (1983) account for these two different types of ironic meaning.

Nevertheless, despite evident issues in categorizing different forms of irony, it is important to remember that, as Kreuz (2000, p. 104) indicates, "the job of the listener is to recover the discourse goals of the speaker and not to identify some rhetorical label like irony or understatement." In [Chapter 4](#), we will discuss the potential issues with the way Taguchi et al. (2016) define irony and the role of conventionality in irony comprehension.

What will however be of a great value for our study is an analysis of the functions, impact and markers of irony, which we will do after we discuss the mechanisms of production and interpretation of irony.

#### *2.6.2. A literature review of the role of literal meaning in the mechanisms of irony production and comprehension in L1*

Numerous studies have explored how native speakers process irony, focusing on how quickly they read and respond to ironic statements (compared to neutral/literal ones) in order to infer processing difficulty or ease. The first theories of verbal irony (for example Cutler, 1974, or Searle, 1975; Searle & Vanderveken, 2009) were based on its traditional definition (the ironic intent was the opposite of the literal meaning, i.e., meaning reversal). Indeed, Grice's Standard Pragmatic Model (1975, 1989) was based on a three-step process by which the "addressee was obligated to process the expressed (or literal sense) of an utterance" (Katz, 2000, p. 2), before they could reject it (because it violates one of the conversational maxims—usually for irony the Maxim of Quality), and access the ironic meaning. Gibbs summarizes the process this way (1986, p. 3):

- (a) compute the utterance's context-independent, literal meaning;
- (b) decide whether the literal meaning is the speaker's intended meaning;
- (c) if the literal interpretation is inappropriate, compute the non-literal meaning by assuming the opposite of the literal interpretation.

This model has been rejected by both proponents of the direct access model (or "processing-



equivalence hypothesis”<sup>7</sup>), as well as supporters of a different kind of “sequential processing” (Attardo, 2000, p. 797) approach. It is important to note that these irony processing models often stem from different theories of irony which either oppose or expand upon Grice’s traditional view, and impact the way hearers may process irony statements depending on their ability to detect and recognize the dissociative attitude of the speaker. Before discuss processing theories, we will briefly review some of the most prominent theories of irony:

- The **echoic mention** theory or echoic interpretation from Sperber and Wilson (Jorgensen, Miller, & Sperber, 1984; Sperber & Wilson, 1981, 1998; Wilson, 2006, 2013; Wilson & Sperber, 1992, 1998, 2002) is based on the idea that ironic utterances always refer to or echo a previously uttered (or insinuated or imagined) comment. Departing from Grice’s two-stage model, this theory “does not involve a double (literal and figurative) deciphering procedure” (Giora, 1995, p. 240). One criticism of this theory is that it does not account for the possibility of ironic asteism (it focuses only on negative feelings/perception).
- The **pretense** theory from Clark and Gerrig (Clark & Gerrig, 1984, and later refined by Currie, 2008) which states that speakers only “pretend” to be someone else or to have a different opinion.
- The **echoic reminder** theory from Kreuz and Glucksberg (Kreuz, 1996, 2000; Kreuz & Glucksberg, 1989) expands upon Sperber and Wilson’s model. It emphasizes the “reminding function of irony based on the shared attitudes and expectations held by conversational participants.” (Gibbs & O’Brien, 1991, p. 526). One of the pillars of this theory (which was questioned by other authors later on) was that the echoic reminder/irony needed to target a specific victim.
- The **allusional pretense** theory from Kumon-Nakamura (Kumon-Nakamura et al., 1995) who criticizes previous theories and claims that irony stems from an allusion to a “failed expectation” (which violates the pragmatic rules of discourse).<sup>8</sup>
- The **indirect negation** view of irony from Giora (1995) (see also Giora, Fein, & Schwartz, 1998) as a “form of negation that does not make use of an explicit negative

---

<sup>7</sup> Giora, Fein and Schwartz, 1998, p. 84 (note that Giora and colleagues are NOT proponent of this approach).

<sup>8</sup> Kumon-Nakamura, Glucksberg, and Brown (1995, p. 3): “The allusional pretense theory claims that ironic remarks have their effects by alluding to a failed expectation. In normal conversation, this is accomplished by violating pragmatic rules of discourse, usually the maxim of sincerity. Such violations simultaneously draw a listener’s attention to the failed expectation and express the speaker’s attitude (normally but not necessarily negative) toward the failed expectation.”

marker” (Giora, 1995, p. 240). Contrary to the traditional view of irony that requires the literal meaning to be rejected, Giora argues that irony as indirect negation retains “both the explicit and implicated messages so that the dissimilarity between them may be computed” (p. 241—this is the basis for her Graded Salience Hypothesis which we will discuss shortly). For Giora, the literal meaning (“marked utterance,” Giora, 1995, p. 245) refers to an “expected/desired state of affairs” (1995, p. 245).

- The **contextual-appropriateness theory** of irony from Attardo (2000a) is an expansion of Grice’s CP. Attardo presents the following conditions that allow for an utterance (*u*) to be ironical (2000a, p. 817):
  1. *u* is contextually inappropriate,
  2. *u* is (at the same time) relevant,
  3. *u* is construed as having been uttered intentionally and with awareness of the contextual inappropriate by S [speaker], and
  4. S intends that (part of) his/her audience recognize points 1-3,
  5. unless H [hearer] construes *u* as being unintentional irony, in which case 3-4 do not apply.

Attardo concludes that in addition to irony being a “purely pragmatic phenomenon” (p. 823) that requires inferential processing, it is “essentially an inappropriate utterance which is nonetheless relevant to the context” (p. 823).

Along with these diverging views on irony, a long-standing debate—usually centering on the role of literal meaning in indirect meaning comprehension—still divides researchers into two indirect meaning processing approaches:

- 1) The **one-stage processing** approaches (also referred to as direct access model or processing-equivalence hypothesis), supported by a series of experimental studies led by Gibbs (1979, 1980, 1983, 1986, 1993, 1994, 2000, 2002) and colleagues (Gibbs & Colston, 2012, Gibbs & O’Brien, 1991, Gibbs, O’Brien, & Doolittle, 1995), Sperber and Wilson (1981, 1986, 1995, 1998, 2002 to only reference a few) and Ivanko and Pexman (2003), adhere to a **parallel processing model** which predicts that, when sufficiently contextually-primed, indirect meanings can be arrived at directly (Bromberek-Dyzman, 2015; Gibbs, 1986). These approaches focus on the role of irony-supportive context, i.e., “comprehension speed is cued by and depends on the degree of context supportiveness” (Bromberek-Dyzman, 2015, p. 217). Gibbs’ experimental studies (1979, 1980, 1986, etc.) rejected the two-stage processing approach (that requires literal meaning to be processed first) by showing that people do not take longer to process non-literal meaning (indirect

requests or idiomatic expressions) than they do literal sentences. According to Gibbs, decoding and rejecting the literal interpretation would have increased processing and reaction time, which means comprehension of indirect meaning would have taken longer. His studies results showed that ironic comments can be comprehended faster or as fast as their literal equivalents, which led Gibbs to posit that people access the non-literal meaning of the utterances without processing the literal meaning because the context primes the hearer to adopt an “ironic frame of mind.”

Giora (whose hypothesis will be explained under the next point) however argues that Gibbs’ findings come from a flawed test design that relies on literal examples involving “explicit negation (which is more difficult to understand and prohibitive) and is pragmatically inappropriate because it is uninformative” (1995, p. 254).

- 2) The **multi-stage or modular approaches**: usually a two-stage process during which the literal meaning is first accessed before the non-literal or ironic interpretation can be reached (Giora, 1995, 2003; Giora & Fein, 1999; Giora, Fein, Laadan, Wolfson, Zeituny, Kidron, Kaufman, Shaham, 2007; Giora, Zaidel, Soroker, Batori, & Kasher, 2000; Filik & Moxey, 2010). These approaches reject the idea that hearers can bypass the literal meaning of indirect utterances. A multi-stage approach usually results in longer processing or response times. This category of processing can be divided into two trends: the traditional approach (cf. Grice) that we mentioned earlier, which requires people to process the literal meaning first and then reject it to access the intended meaning, and approaches that align with Giora’s Graded Salience Hypothesis (GSH). In the latter, the literal meaning is activated first but is not discarded: because Giora sees irony as an indirect negation (see below), she argues that the literal (more salient) meaning is “not suppressed but remains active” (Giora et al., 1998, p. 91) when processing irony. Her theory centers around the idea of salience, which she defines as a function of its conventionality, familiarity, frequency, or “givenness status in a certain (linguistic or nonlinguistic) context” (1997, p. 185). For Attardo, the GSH rejects “the primacy of the literal sense [i.e., the Grice’s Standard Model], in favor of the primacy of the salient sense (be it literal or figurative/indirect)” (2000, p. 801). Giora and Attardo seem to disagree on the role of context in processing and understanding indirect meaning however. Giora believes that “context has a limited role” (1997, p. 186) in processing order because while it may “facilitate activation of a word’s meaning, it does not inhibit activation of the salient (e.g., frequent), meanings” (1997, p. 186). On the other hand, Attardo’s view of irony as **relevant inappropriateness**, emphasizes on the role of co(n)text and suggests that (p. 821) “when we encounter a contextually inappropriate

element in the text, that element becomes highly informative and hence salient.” His analysis of irony processing derives from Grice’s Cooperative Principle to which Attardo adds a fifth Maxim “to be appropriate.” Attardo argues that his relevant inappropriateness theory is not necessarily incompatible with Giora’s GSH. To him, “contextual salience” falls under the GSH’s fourth metric (see above: “givenness status in a linguistic context”) and is related to the Giora’s concept of informativeness (1995, 1997): “the least appropriate element will tend to be the most informative” (Attardo, 2000a, p. 821) which leads Attardo to conclude that “contextual inappropriateness may trigger salience” (2000a, p. 821).

The above theories are very condensed and summarized accounts of different researchers’ views of indirect meaning and irony processing. One thing we found particularly important for our analysis of indirect meaning comprehension is the concept of context-priming. We envision the possibility that salience (whether we are talking about salience of meaning or salience of context) may impact processing times and ease of understanding. For example, following Giora’s GSH, we agree that in situation where the literal meaning of an utterance is particularly salient, the literal meaning will likely be processed first. However, a salient indirect meaning (e.g., a very familiar idiomatic expression that is usually used in its indirect meaning) may be processed first. We believe contextual cues also play an important role in making direct or indirect meaning more salient, and prime the hearer for a specific type of processing. Before we review ways in which access to contextual cues may impact processing of indirect meaning, we will briefly discuss some of the functions of verbal irony. Being able to perceive the attitude of a speaker and reasons as to why they use irony is an important key to understanding irony.

### *2.6.3. Functions of verbal irony*

Based on irony processing theories, some authors, like Dews and Winner (Dews, Kaplan, & Winner, 1995; Dews & Winner, 1995), have shown that the literal meaning of an ironic utterance, whether it is meant as a praise or a criticism, allows to “color” or “nuance” the non-literal intended message conveyed by the utterance. This is what they called the **Tinge Hypothesis**. They explain that an indirect criticism disguised as an apparent praise (saying something positive to express something negative) has a less negative impact compared to a direct criticism (“irony is perceived as muting the underlying evaluative message,” Schwoebel, J., Dews, S., Winner, E., & Srinivas, K., 2000, p. 49). This irony diminishes the level of aggression between the speaker and the target of irony (Brownell, Jacobs, Gardner, &

Gianoulis, 1990; Kreuz, Long, & Church, 1991; Milanowicz, 2013; Myers Roy, 1981). On the other hand, according to Dews and Winner, a compliment given in the form of an ironic comment (disguised as a criticism—and asteism), does not quite reach the level of positivity that a direct compliment would have. This analysis of irony processing implies that one must access the literal meaning of the utterance, even if it is processed in parallel to the ironic meaning processing. Dews and Winner state (1995. p.7): “the literal meaning is processed because it serves the pragmatic function of muting the speaker’s meaning.” Their theory thus contradicts Gibb’s initial hypothesis that literal meaning does not have to be considered to reach the ironic meaning. In that aspect, Dews and Winner concur with Giora’s **indirect negation** theory (Giora, 1995; Giora & Schwartz, 1998), according to which the literal meaning must be activated to make sense of a negative ironic statement.<sup>9</sup>

Some authors, such as Attardo (2000a, 2002, 2017a, 2017b et Attardo & Raskin, 2017) gave us valuable insights into the functions of verbal irony. Among the other reasons why someone might use irony, we can add the following (Attardo compiled a list of them based on previous research from various authors):

- *Group affiliation:*
  - The speaker shows off his detachment and superiority/wit/ability to play with language (Haiman, 1989)
  - Inclusion: the purpose of the irony is to bond. It reinforces in-group solidarity and complicity. (Lakoff, 1990; Myers-Roy, 1981)
  - Exclusion: mockery and sarcastic criticism used to make fun or criticize someone outside of the group. (Myers-Roy, 1981)
- *Entertainment/fun* (Shively, 2013)
- *Sophistication:* showing mental dexterity, wit or to be funny (Dews et al., 1995; Kreuz et al., 1991)
- *Retractability:* irony is a face-saving device. It is non-committal (Attardo, 2002) and its muting functions allow to preserve relationships (cf. face-saving device in Che, 1990, Dews et al., 1995, and Jorgensen, 1996). Berrendonner (1981) also explains that it allows the speaker to not be held accountable and to contradict themselves, saying they

---

<sup>9</sup> Giora and Schwartz (1998, p. 83): “According to the indirect negation view of irony (R. Giora, 1995), irony comprehension should involve retention of the activated literal meaning so that the comprehender may compute the difference between the (usually desirable) state of affairs alluded to by the literal meaning and the less desirable, ironized situation. Evidence from three experiments supports the graded salience hypothesis and the indirect negation view of irony.”

were just joking, to avoid any possible negative consequences of sanctions.

- *Nuance of Evaluation*: Based on Dews and Winner's Tinge Hypothesis (1995), we know irony allows a person to express various degrees of criticism (less harsh) or praise (less positive).
- *Persuasive, memorable and informative aspects* (cf. Carston, 1988; Giora, 1995; Kreuz et al, 1991; who all agree that irony is a powerful rhetorical device that can persuade, ensures people will not forget).
- *Politeness strategies* (in particular when someone is expressing something negative. Cf. Barbe, 1995; Brown and Levinson, 1978; Chen, 1990; Dews & Winner, 1995; Giora, 1995). As mentioned previously (2.4), irony can be a face-saving device that allows the speaker to appear less rude or threatening (Brown and Levinson, 1978; Jorgensen, 1996; Gibbs, 2000, Boxer, 2010; Colston, 2017 to only cite a few).
- *Silence breaker/filler* (Shively, 2013, whose experiments analyzed the functions of humorous utterances, including ironic utterances).
- *Deflecting a compliment* (Ishara & Cohen, 2010; Shively, 2013)

Now that we have reviewed the different functions of irony, we will turn our attention towards the markers or cues that can help detect verbal irony. As previously stated, we believe that contextual cues play a very important role in irony comprehension and several authors (particularly Attardo and Yus) have shown how these markers or cues are not only devices on which speakers rely to produce irony, but also clues to be able to detect and interpret it.

#### 2.6.4. *Accessing the ironic interpretation*

Following Attardo's theory of relevant inappropriateness, Yus (1998, 2000, 2017) argues that the activation of contextual sources and the detection of incompatibility or incongruency with the literal meaning (what is said) will enable access to the ironic interpretation of utterances. The more numerous, better, or more salient these incompatibilities are, the easier it will be for people to detect the irony (Yus, 2000, p. 15):

The processing effort required for the interpretation of the intended ironic meaning of an utterance decreases in proportion to the increase in the number (and quality) of incompatibilities (detected by the addressee) between the information supplied by the inferential integration of simultaneously activated contextual sources (leading or leading plus supportive) and the information provided by the proposition expressed by the utterance.

These criteria for **optimal accessibility** to irony are based on the relevance theory account of irony from Sperber and Wilson. In order to recover the ironic meaning of an utterance, the hearer must identify the echoic quality of that utterance as well as the “speaker’s attitude of dissociation towards the utterance” (2000, p. 5). While Yus explains that the activation of multiple cues often occurs, there is no way to predict which source will be accessible and activated and how strong the activation will be. This concurs with what authors have said about irony markers: “irony markers cannot be defined as infallible pointers to irony” (Muecke, 1978, p. 365). An irony indicator can alert the audience or reader to the fact that a comment is ironical but the utterance/sentence could still be ironical without this marker. Below is a list of seven irony-relevant contextual sources (some of which rely on these irony markers), that form the basis of Yus’s optimal accessibility (presented in Yus, 1998 and developed in Yus, 2000, pp. 5-12)—we should note here that we believe most of the contextual sources below will also apply to other types of indirect meaning:

- **Encyclopedic or factual information:**

Following the Relevance theory, people usually use mental representations and stereotypical information as a “mental background against which all new incoming information is processed” (also cf. Sperber & Wilson, 1986). (this includes mental representation, stereotypes, societal norms, ideals, expectations, common sense assumption, etc.). Yus classifies them into four categories:

- 1) **Macrosocial norms and factual information:** factual knowledge and assumptions, or “prior cognition” as Hamamoto (1998) calls it, are used as a general background to all forms of language processing, but Barbe (1989, 1995) stipulates that irony requires “shared general knowledge among the participants, in-group as well as societal or cultural knowledge.” Indeed, irony users often rely on cultural norms to build their irony strategies and hearers must be able to detect the incompatibility between these cultural norms or prior knowledge to detect the contradiction or contrast created by the ironic statement.
- 2) **Commonsense assumptions** are frames or schemas that people use to process prototypical conversational situations, and particularly for irony recognition. Sperber and Wilson (1995) provide an example of how this contextual source can be used to make sense of the following sentence: “The child left the straw in the glass.” In this example, the hearer will understand that the speaker is referring to a drinking tube (and not a cereal stalk) when using the word “straw” because this

interpretation (which is reached unconsciously and with minimal processing effort) makes sense given the particular conversation situation. Kreuz and Roberts (1995) also explain that hyperbole and understatement, two forms of ironic strategies, rely on these types of shared stereotypical values. This process of commonsense assumption is also illustrated by Giora and Fein (1999) with the following example (cited in Yus, 2000, p. 6):

Only by assuming a default, stereotypical idea of the amount of food that one can eat at a time, can the understatement-based irony (11b) in situation (11a) be successful (cited in Giora and Fein, 1998). The mother echoes a parallel situation such as (11a), a situation in which she could have uttered (11b) descriptively (and non-ironically):

- (11) a. [*After he has finished eating pizza, falafel, ice-cream, wafer and half of the cream cake, Moshe started eating coated peanuts*].  
b. Mother: “Moshe, I think you should eat something.”  
c. [*Moshe has not eaten anything for several hours*].

3) **General/prototypical expectations** also apply to how conversations usually progress: as such, Haverkate (1990) explains that irony violates the felicity conditions that govern conversations and Alba Juez (1998) provides a list of “conventionalized strategies” by which irony violates the principle of collaboration of on-going conversations (cited in Yus, 2000, p. 6):

- “answer an obvious question with an even more obvious question”—the Pope question.
- “reply to a lie with an even bigger lie” —an adynation.
- “reply to a stupid question with an even more stupid answer”
- “ask a question and give a ridiculous answer before the hearer can answer himself/herself.”

4) **Microsocial, situational expectations** are contextual information or facts known by the speaker and hearer. These are often the “only reliable source for the identification of irony” (Yus, 2000, p. 7). Unlike the first three types of prior or shared knowledge, these do not rely on cultural norms or factual information but rather are very specific to the particular situation the speaker/hearer found themselves in. For example, if a teacher were to say “This is a long paper!” to a student who just turned in a 3-page paper, when the teacher had previously stipulated that term paper should be around fifty pages, that prior knowledge would be necessary for the student to understand the ironic comment. Similarly, if the student hands in a 300-page paper and the teacher exclaims “This is a very short paper!,”



the student would rely on the same situational expectation (i.e., the paper was supposed to be around 50 pages) to evaluate the mismatch between situations and utterances and interpret the comment as ironic (see Glucksberg, 1995, p. 53). The speaker (teacher) echoes potential utterances that may have been true in a different situation.

- **Mutually manifest physical environment (setting):**

The physical context around people engaged in conversation is part of the participants' cognitive environment during their interaction. Through perception of their surroundings, both speaker and hearer receive input which is used to build assumptions (through the "echoic quality of the utterance and the speaker's attitude of dissociation towards this utterance." Yus, 2000, p. 30), which can then be utilized for irony. For example, uttering "Lovely weather today!" when it is in fact pouring rain and both speaker and hearer can assess the outside weather conditions, relies on this type of "mutually manifest physical environment source" to be interpreted as ironic. This also fits with what Barbe calls "situational irony," which focuses on "participants' knowledge about a particular situation, especially immediate situation (Barbe, 1989, p. 277).

- **Speaker's nonverbal behavior:**

Yus includes both kinesic and paralinguistic cues in this category. According to him, these tools can be used by speakers to "facilitate the identification of an ironic message" (Yus, 2000, p. 8) as well as indicators on which hearers can rely to detect the irony and interpret the intended message. Kinesics markers include gestures, facial expressions, like winks, nudges, raised eyebrows, or "blank face" (Attardo, Pickering, & Baker, 2013; Attardo & Poggi, 2003; Guidi, 2017; Haugh, 2017; Jobert & Sorlin, 2018; Wallmark, 1971; Yus, 2000, 2017).

From a phonological perspective, opinions differ regarding the existence of an "ironic tone" or rather, the fact that there may not be a single specific intonation pattern that can be attributed to irony, but rather a plethora of phonological devices, or "contrasting phonetic features" (Bryant, 2013, p. 117) and "vocal signals" (Bryant, 2013, p. 104) that can be used when conveying an ironic message (Bertrand & Prievo-Valverde, 2013; D'Imperio et al., 2013; Gibbs, 2000; Gibbs & O'Brien, 1991; Wallmark, 1971; Yus, 2000;) such as high-pitch/rising intonation (Schaffer, 1982), contrastive-exaggerated intonational patterns (Cutler, 1974; Barbe, 1995; Schaffer, 1982), or contour (Shapely 1987), lower pitch, falsetto, singsong melody for sarcasm (Haiman, 1998; Muecke, 1978), softening of the voice (Muecke, 1978), nasalization (Chen, 1990; Cutler, 1974; Haiman, 1998; Schaffer,

1982), differentiated stress pattern, vowel hyper-articulation, etc. (Attardo et al., 2013, Ivanko & Pexman, 2003; Wallmark, 1971). While some authors (Clark & Gerrig, 1984; Cutler, 1974, 1976; Kreuz & Glucksberg, 1989; Sperber & Wilson, 1981, 1998) do refer to an ironic tone (heavy stress, slow speaking rate, nasalization), it is often in reference to the pretense, mention or echoic reminder theories of irony they support. Indeed, in Clark and Gerrig's (1984) pretense theory for example, the "ironist is pretending to be an injudicious person speaking to a naïve audience" (Kreuz & Roberts, 1995, p. 22) and as such the speaker may use this "ironic tone" to pretend to be someone else. Ackerman (1983) also shows that an "ironic tone of voice" could help children distinguish between irony and deceit. However, this author also found that contextual cues were more important than intonation in children's understanding of ironic intent. A theory that was also supported by Winner and Gallagher (1983—cited in Kreuz & Roberts, 1995) who found that behavioral cues (like pointing or laughing) were more informative than intonational cues in adults. Therefore, while non-verbal cues may be important, "if cues from the context are strong enough, no intentional cues are necessary at all" (Cutler, 1974, p. 117). Kreuz and Roberts (1995) also point out that the ironic tone may be superfluous or even detrimental since "an important goal of irony is to be humorous" (Roberts & Kreuz, 1994, pp. 28-29). If an ironic statement is made too obvious, then it may "lose its ability to amuse and may elicit emotions of a very different kind" (p. 28-29).

- **Mutual knowledge and addressee's familiarity with the addresser:**

Yus (2000) differentiates between what he calls the "addressee's background knowledge of the addresser's biographical data" (pp. 8-9) and "mutual knowledge" (pp.9-10), which he presents as two separate sources. The former are the beliefs or assumptions the hearers and speakers may have the other person's opinions, preferences, etc. as well as an "awareness of the degree of familiarity which relates both interlocutors" (Yus, 2000, p. 9). The latter (mutual knowledge) is information that interlocutors assume they share, what Sperber and Wilson's (1986) refer to as "mutually manifest assumptions" which may overlap with other contextual sources. To this we can also add the personality of the speaker and hearer and the ability for the hearer to recognize the intent of the speaker based on this personality (Wallmark, 1971). The following example is used by several authors (Dews & Winner, 1995; Gibbs, 1994; Haverkate, 1990): "John is a real Einstein." Here, the speaker must assume that, based on their specific context, the hearer's opinion of John's intelligence is already low, and that the hearer will use this mutually manifest assumption to recognize the fact that the speaker is being ironic (dissociative attitude towards the expressed

proposition—cf. Yus, 2000). This also aligns with what Barbe (1989) refers to as “personal irony” as it relies on the speaker’s “knowledge of the interlocutor’s cognitive environment” (Sperber & Wilson, 1986, p. 45). Mutual knowledge can also be constructed during the conversation so it is not static and the speakers are “constantly checking the status of this information that they supposedly share with their interlocutors” (Yus, 2000, p. 9). The hearer’s awareness of this mutual knowledge is key to the process of identifying the speaker’s attitude of dissociation (i.e., the ironic intent).

- **Role of previous utterances in the conversation:**

Being able to refer back to previous utterances (and their interpretation) to understand the context of an ironic statement can be vital to the decoding of the intended message of the speaker. This is particularly true when the ironist bases his irony strategy on the repetition of a previously uttered statement. A clear example of this can be found in Sperber and Wilson’s echoic mention theory (1981) and Cutler’s provoked irony (1974). Cutler provides the following examples where “the speaker repeats back to his audience something which the audience has previously said or heard, with the ironic intonation indicating that the previous claim has turned out to be wrong” (Cutler, 1974, p. 118-119):

(a) *Sure, Joe, you locked the door.*

(b) *At least it won’t rain, he says.*

(c) *The cops won’t give us any trouble, Harry’ll handle them.*

It is easy to imagine that (a) echoes Joe’s previous words about locking the door, and indicates that he failed to do so. In (b) the speaker repeats what someone has said regarding the weather condition, which the hearer would have also heard. A similar echoic mention can be found in example (c) where an indirect quotation of what Harry likely said previously (that he will handle the cops and that the cops will not give them any trouble) is being reused sarcastically (implying that Harry is not handling the issue). Yus (2000) illustrates the role of previous utterances with an example from Winner (1997), in which a mother ironically says “I see you are wearing your sweater,” to her son who is not wearing a sweater even though she has just asked him to wear one.<sup>10</sup> Of course, this will remind us of the situational expectation sub-category listed previously—clearly, contextual sources can belong in more than one category, and we know that a specific situation could also activate multiple sources at the same time. In the above example, Yus also points out the following (2000, p. 10):

---

<sup>10</sup> The role of the context has also been tackled by researchers dealing with written format. Indeed incompatibility detected between the elements of the same sentence or larger text could cue the reader in as to the presence of irony (cf. Kerbrat-Orecchioni, 1976, or Muecke, 1978)

The mutually manifest visual evidence of the son not wearing [a sweater], together with the reference to that previous request, are easy-to-access interpretive aids foregrounding the mother's ironic attitude of dissociation towards her ironic statement, and at the same time she ironically underlines (i.e., makes mutually manifest) the existence of a more desirable potential situation [the son is wearing his sweater] in which she would have endorsed [the statement she made as a literal comment "I see you are wearing your sweater"].

- **Linguistic cues:**

While some other authors (Barbe, 1995; Myer Roy, 1981; Wallmark 1971; Yus, 2000) agree that linguistic devices (lexical, syntactic, stylistic, etc.) are not irony-specific (since the same devices can be used for literal meaning), they state that they may "alert" the hearer to the possible ironic quality of the utterance (Yus, 2000, p. 10). We have already discussed the phonological or prosodic devices (see "Speaker's nonverbal behavior above"), which Seto (1998) and Wallmark (1971) list under linguistic devices, but drawing from various authors, we will present the following list of additional linguistic cues:

- 1) ***Lexical devices:***

- Overstatement (lexical words showing exaggeration: e.g., really, extremely, quite a bit, etc., Kreuz et al., 1991; Kreuz & Roberts, 1995; Seto, 1998; Wallmark, 1971), understatement (lexical words weakening a situation: e.g., sort of, a little, not, etc., Wallmark, 1971), "outmoded words" (e.g., "Fat chance," "That's a likely story") or "sarcastic idioms" (Barbe, 1993—use of "isn't it ironic that..."). Kreuz and Roberts (1995, p. 24) also point out that the combination of adverbs and extremely positive adjectives can be used to convey an ironic message.

- 2) ***Syntactic devices:***

- Exclamations (idioms and interjections, Seto 1998; Wallmark, 1971), Yes/No questions, rhetorical questions (Pope question), echo questions, commands (Kreuz et al., 1991; Wallmark, 1971), superlative and focus topicalization (Seto, 1998).

- 3) ***Stylistic devices:***

- Politeness (Wallmark, 1971) and honorific (Seto, 1998), or as Glucksberg (1995, p. 52) puts it, "strikingly overpolite constructions" (such as thanking someone after they step on your toes, or offering more food to someone who just ate a lot).

- 4) ***Other tropes*** (which may require a combination of linguistic devices):

- Adynaton (an extreme form of hyperbole), oxymoron (use of apparently contradictory terms), hyperbole (exaggerated statements), metaphor, meiosis (understatement), litote (expressing an affirmative by the negative; e.g., “he is not a bad singer” to say someone is a good singer), (Gibbs, 2000; Seto, 1998), euphemism, repetition (Yamanashi, 1998), or “conventional set phrases” (Kaufer, 1981; Sperber & Wilson, 1981) such as catch phrases, slogans or maxims.

While we cannot imagine that all different types of irony markers (phonological, kinesic, semantic, etc.) would be used at the same time, it is not uncommon for ironists to rely on more than one. Similarly, more than one type of contextual sources may be activated during the processing of an ironic statement. Yus (2000) argues that one contextual source will be more accessible or salient: he labels this the *leading contextual source*. All other sources that may be activated during irony comprehension are what he calls *supportive contextual sources*. Summarizing his criterion of optimal accessibility to irony, Yus (2000, p. 15) maintains that:

The processing effort required for the interpretation of the intended ironic meaning of an utterance decreases in proportion to the increase in the number (and quality) of incompatibilities (detected by the addressee) between the information supplied by the inferential integration of simultaneously activated contextual sources (leading or leading plus supportive) and the information provided by the proposition expressed by the utterance.

Therefore, when presented with a highly salient leading contextual source, the hearer will detect the irony, and the activation of supportive sources then allow for a decreased processing time and effort.

Another important source of irony processing that also highlights the importance of context is the hearer’s knowledge or perception of the speaker’s communicative style or preferences. Yus’ addressee’s background knowledge of the addresser’s biographical data did not quite address the possibility that the hearer may take the speaker’s tendency to use irony into account to activate predictions and processing of irony. Regel, Gunter, and Friederici (2010) conducted two experiments to see how the frequent use of irony by a speaker may trigger the anticipation of irony in the hearer. They found that when a speaker displays preferences for irony comments and a more implicit style of communication (using indirect meaning like irony instead of explicitly stating their point), the hearer starts expecting irony and is able to process and detect it. According to Regel et al., it is that expectation that impacts the irony processing pattern more

so than the literal/non-literal dichotomy of the message. Similarly, Ivanko and Pexman (2003) show that certain contexts may trigger different expectations of irony use. For example, strongly negative situations will often lead the hearer to expect literal criticism while a more moderately negative situation will prompt the hearer to anticipate possible irony, thus facilitating the processing of the ironic criticism.

As mentioned previously, we believe that Yus's theory of optimal accessibility and accompanied contextual sources can also be applied to other types of indirect meaning as the tenants of this theory revolve around the ability to detect an incompatibility or incongruity between what is said and the general context or specific situation. In this section we have discuss the mechanisms of irony production and interpretation; we will now turn our attention to one final field of linguistics which may shed some light on the other two types of indirect meaning we are focusing on in this dissertation. Indeed, Conversation Analysis (CA) provides some insightful perspectives in how interactive patterns of communication can predict or prime certain indirect meaning interpretation.

## 2.7. Indirect meaning from the perspective of Conversation Analysis.

One major difference we would like to point out between irony (which we just discussed extensively in the previous section) and the other two types of indirect meaning this dissertation is focusing on (i.e., indirect refusals and indirect opinions) is the fact that refusals and opinions can also be expressed in a direct manner (the same claim could probably not be made about "indirect or direct irony"). One approach that gives us useful tools to analyze patterns of conversation that include refusals and opinions is Conversation Analysis (CA). CA provides a way to analyze and study the pragmatics of everyday discourse and communicative exchanges. It draws from the previously reviewed theoretical framework (Cooperative Principle, Relevance Theory, Austin and Searle's Speech Acts, and even more specifically the theories of Politeness we previously reviewed) and puts a specific emphasis on meaning negotiation and turn-taking patterns. In this section, we will focus specifically on indirect refusals and indirect opinions and discuss the potential perspective that CA brings in understanding the mechanisms behind comprehending these two types of indirect meaning.

Two particular aspects of CA that we wanted to discuss in this section (because of their

importance for indirect refusals and indirect opinions) are sequence organization and the concept of preference.

### *2.7.1. Sequence organization and turn-taking*

Schegloff (2007) defines “sequence organization” as the “organization of courses of action enacted through turns-at-talk – coherent, orderly, meaningful successions or “sequences” of actions or “moves.” Sequences are the vehicle for getting some activity accomplished” (p. 2). For example, when two people meet, they often exchange greetings in a sequence: one person will greet the other who may in turn respond with a similar action (possibly another form of greeting). These basic pairs of conventionally related actions are referred to as “adjacency pairs.” This is an expected conventional sequence common in many languages. This pattern of interaction is important because it creates an expectation that can help interpret indirect meaning.

The first and second part of the adjacency pair unit may call for different types of actions: an invitation (FPP: first pair part) may lead to an acceptance (SPP: second pair part), an offer (FPP) may be met with a refusal (SPP). This conventional pattern creates an expectation of what the SPP would be. Thus, in the case of refusals, whether the FPP presents as an offer, request, invitation, suggestion, or advice, the adjacency pair rule puts a turn-taking constraint on the SPP which is expected to represent an act of acceptance or refusal (Chen,1996; Kerbrat-Orecchioni, 1992, 2008; Manno, 1999; 2002; Mbow, 2011; Mulo Farenkia, 2015, 2020; Sacks, Schegloff, & Jefferson, 1994; Schegloff, 2007; Schegloff & Sacks, 1994; Taguchi, 2007, 2008b, 2008c).

When it comes to expressing solicited opinions, we often see a similar pattern of interaction. As Taguchi (2002) highlights “a question asking for evaluated comments should be followed by “good” or “bad.” When asked for information the person usually provides the requested information” (p. 166). Of course, that does not mean that conversation always follow the adjacency-pair rule. If the first person’s FPP is a request for an opinion, the second person may opt out and break the pattern of conversation through different strategies (avoidance, change of subject, explicitly stating they do not want to share their opinion, etc.). On the other hand, one may provide their unsolicited opinion when they were never asked for it in the first place. Yet, as far as CA is concerned, the adjacency pair rule places an expectation on the conversational

exchange such that if the FPP includes a request for evaluative judgment, we will assume the SPP to be a statement of opinion. Which means that interlocutors will process the SPP through the lens of that expectation, whether the opinion is expressed directly or indirectly. As Taguchi states—referencing Hutchby and Drew’s (1995) work on Conversation Analysis—“responses that diverge from the expectation of the adjacency pair framework tend to be marked with intonation, pause, or hesitation markers” (Taguchi, 2002, p. 166).

### 2.7.2. Preference organization

In addition to the sequence organization (and particularly the adjacency pair) priming the interlocutor for a specific pattern of turn-taking exchange, the term preference is also used to refer to a specific type of SPP that is most commonly used and expected in conversation. Brown and Levinson (1978) define preference as follows (p. 38)

The term refers to the phenomenon that after specific kinds of conversational turn, responses are often strictly non-equivalent: one kind of response, termed the *preferred*, is direct, often abbreviated and structurally simple, and typically immediate; in contrast, other kinds, termed *dispreferred* are typically indirect, structurally elaborated, and delayed (Pomerantz 1975, 1978, 1984a). The preferred type of response is usually more frequent also, but the term ‘preference’ refers to the structural disposition, to the fact that conversational organization conspires to make it easier to use the preferred type of turn, not to participants’ wishes.

They provide examples of such preference patterns: the preferred turn for an offer would be an acceptance, the dispreferred turn a rejection, the preferred for a question would be an answer (the dispreferred turn would be a non-answer or lack of response), etc. Of course, we already know from our review of the Politeness theories ([section 2.4](#)) that depending on the context, a preferred turn could result in a more face-threatening act. In the case of an invitation, acceptance (the preferred turn) would be seen a less face-threatening than a refusal (dispreferred turn). However, for opinion requests, it would depend on whether the evaluation is positive or negative. Indeed, if one asks an opinion, the preferred turn would be for the SPP to provide the requested opinion. If the opinion is a positive, favorable evaluation, the SPP will not be considered a face-threatening act. However, if the requested opinion is one that is negative, the preferred turn can then threaten the interlocutor’s face. This is where indirectness and politeness come into play.

We know that direct (explicit) refusals can be realized in many different ways by using



illocutionary force indicating devices (IFIDs) such as “no,” “I cannot,” which Mulo Farenkia (2019) label “negative ability realization” (p. 582) (referencing Beebe et al. 1990). However, direct refusals are considered a dispreferred turn and present as a face-threatening act. In his analysis of French people’s refusal production, Mulo Farenkia (2015, 2019, 2020) noted that French people will often mitigate their direct refusals with expressions of gratitude (“*Non merci*”) following by a justification for their refusal and/or suggestion. These additions are meant to “soften the illocutionary force of the face-threatening act” (Mulo Farenkia, 2015, p. 582) that is refusal (cf. Theories of Politeness in [section 2.4](#)). They are often referred to as adjunct refusals, and several authors have observed these strategies used in languages other than French: English (Taguchi, 2002, 2005, 2008a, 2008c, etc.; Mulo Farenkia 2019, 2020; Tanck, 2002 to only name a few), Korean and Norwegian (Krulatz & Dixon, 2020), Chinese (Chen, 1996; Ren, 2015), Arabic and Hebrew (Stavans & Shafran, 2018). Finally, indirect refusals can take several forms as well: justification, apology, promise, reproach, etc., and are different from the previous two categories (direct and adjuncts) in the sense that they do not explicitly use an illocutionary force indicative device<sup>11</sup> in the SPP. Overall, authors agree that, while French people may be known to be more direct than other cultures, they favor indirect and adjunct refusals over direct refusals. Yet, the actual realization of the indirect or adjunct refusals may also take different forms depending on the context, power and distance relations (cf. Brown & Levinson, 1978).

Just as Conversation Analysis tools can be useful for understanding the use of indirect refusals, the adjacency pair rule and turn preference can also shed light on the interpretation of indirect opinions. Indeed, it sets the expectation that if an opinion is requested, the interlocutor will provide it. However, a negative opinion may take the form of an indirect speech act in order to soften the illocutionary force of that face-threatening act. Indirect opinions may not follow linguistic conventions or formulas but due to the adjacency pair and preference organization of conversation, a certain pattern of interaction is still present.

---

<sup>11</sup> It should be noted here that IFIDs can include facial expressions, gestures, intonation, etc. However, when we are talking about indirect refusals or indirect opinions for the purpose of this dissertation we are referring to refusals or opinions that are not explicitly “said” (cf. difference between explicature and implicature in section 2.2 and 2.5).

## Background conclusion

The theoretical framework we have presented thus far helps us understand the process by which a hearer may understand indirect meaning by shedding some light on the comprehension processes involved in the interpretation of implicatures. As mentioned previously, a lot of research has been done on irony comprehension but we believe the multi-stage processes described in the sub-section above can also apply to other types of indirect meaning. The fact that hearers may have access and activate multiple cues to comprehend non-literal meaning can be said about indirect opinions and indirect refusals as well as irony. Indeed, gestures, facial expressions, intonation, and of course contextual cues can also be present with any types of indirect meaning and may help the hearer detect the indirectness of an utterance. Supported by the concept of leading contextual source from Yus (1998, 2000), we believe that when faced with indirectness, hearers will try to find relevance in the speaker's utterance by attending to certain cues, but it does not necessarily mean that all cues will be processed and activated. The search for cues may depend on the level of transparency or opaqueness of the indirect utterance (Thomas, 1995; Weizman, 1985), but the number of cues that will eventually be processed may depend on their saliency. A factor that may influence the transparency of an utterance at the sentence-level would be conventionality. We mentioned earlier the fact that "conventionality" has been defined in multiple (and sometimes problematic) ways, but overall, many authors believe that the use of conventional linguistic expressions, or whether a communicative exchange follows a conventional pattern of interaction will make the indirect meaning easier to process (and therefore will not require the hearer to attend to other/as many cues, thus reducing the processing effort—cf. Wilson & Sperber, 2007).

In our next chapter ([Chapter 3](#)), we will review some of the literature on L2 indirect comprehension as it relates to the variables present in our study: proficiency and amount of instruction or exposure, types of implicature and conventionality, and inferencing strategies. Section [3.1](#) will discuss studies that investigate the impact of proficiency on the comprehension of indirect meaning. In section [3.2](#) we will focus on the different types of indirect meaning and how conventionality or opaqueness may affect learners' comprehension according to various authors. Finally, we will look into studies that investigated the use of different strategies by learners in order to interpret conversational implicatures (section [3.3](#)). We should note here that other factors that may impact L2 comprehension of indirect meaning have been studied by L2

researchers, such as learning environment (foreign vs. second language or study abroad experience—cf. Bardovi-Harlig & Bastos, 2011; Shively, 2011, 2013), cultural backgrounds or native language (Bouton, 1988), or individual differences (Kuriscak, 2010). A number of papers and treatment studies also investigate the teachability of implicatures (such as Bardovi-Harlig & Vellenga, 2012; Bouton, 1990, 1999; Eslami, 2010, Ishara & Cohen, 2010, Kim & Lantolf, 2016, Kubota, 1995; Martinez-Flor & Uso-Juan, 2010; Rose, 2005). However, these factors are not directly related to our study and therefore will not be included into our literature review.

## Chapter 3 - L2 Indirect meaning comprehension: a literature review

Many studies have shown that foreign language learners have a tendency to interpret utterances literally, especially at the beginning of the learning process when proficiency is low (Bardovi-Harlig & Dörnyei, 1998; Bardovi-Harlig, 1999, 2013, 2014; Rose, 2005; Rose & Kasper, 2001 ; VanPatten, 2004). From a developmental standpoint, linguistic competences (semantic, grammatical) are often favored in the beginning and the cognitive effort that is required for the decoding of basic elements (sometimes word for word) usually leaves little room for the pragmatic treatment of utterances (which is needed for indirect meaning comprehension). Taguchi (2008c, p. 425) says that “to become pragmatically competent, one needs to be linguistically competent and able to both access contextual information and use linguistic resources appropriately according to context.” Yet, several studies have also demonstrated that even for more advanced learners, figurative or non-literal interpretations—for example conversational implicatures in Bouton’s (1990, 1994a, 1994b) papers, or conventional expressions<sup>12</sup> in Bardovi-Harlig (2009), or Bardovi-Harlig and Bastos (2011)—remain difficult. Development of the pragmatic competence (which is necessary for understanding indirect meaning) is an essential element of foreign language learning, one that is sometimes neglected in order to focus on teaching linguistic competences (semantic, syntactic, etc.). Pragmatic competence development is particularly important because we know that communication often goes beyond what is said (whether speakers use conversational implicatures to convey meaning indirectly, or whether interpretation of their utterance relies on sociopragmatic rules that are relevant to a specific culture). Native speakers rely on pragmatic skills (such as inferencing strategies to understand indirect meaning) to communicate all the time. Bouton highlights both the importance and commonality of conversational implicature inferencing strategy as an **“absolutely unremarkable and ordinary conversational strategy”** [quoting Green, 1989— p. 89], which should be “very much part of any proficient speaker’s communicative competence” (Bouton, 1994b, p. 89).

Different factors (proficiency level, type of indirect meaning, conventionality, cultural knowledge, learning environment, etc.) have been studied by L2 researchers for their potential impact on learners’ comprehension of indirect meaning (whether they may affect accuracy and speed of comprehension, and influence the type of inferencing strategies learners use). [Chapter 3](#) will review findings in the existing L2 literature as they relate to these different factors and

---

<sup>12</sup> We should note that conventional expressions are not necessarily figurative or non-literal

strategies. It should be noted that some of these factors interact with one another in the way they impact comprehension or strategy selections. Additionally, we purposely exclude Taguchi et al.'s (2016) paper from the current chapter as a specific chapter ([Chapter 4](#)) will be devoted to an analysis of their study, which we are replicating.

As a point of reference, the table below presents an overview of the major studies we reviewed for this dissertation whose findings are directly linked to our research focus. We included the following information: author(s) and publication year, target language, participants, test design and measure instrument, types of implicatures or speech acts studied, and research focus. For the participant, we provided the number (if known, the breakdown by participant groups), native language, level of proficiency or amount of instruction/exposure, as well as educational setting when available. In some cases, participants in the control group (native speakers) were not the focus of the study, but data from the control group was analyzed so we also provided some information about the control group in our table. The Test Design/Format/Measure column summarizes information regarding the type of study (longitudinal, follow-up, treatment study, etc.), the type of tests used to collect data (proficiency tests, written or listening tests, multiple-choice questionnaire (hereinafter MCQ), etc.) and any additional details we deemed relevant and important (e.g., number of tasks, whether distractors/fillers were used, etc.). For the types of implicatures or speech acts used, we included specific categories when provided (e.g., refusal to requests, refusals to invitations, refusals to suggestions) along with number for each item when known. For papers that classified the items based on conventionality, we also included that information in the table. Finally, the last column briefly presents the research focus of the papers: the variables the researchers studied (e.g., proficiency level, type of implicatures), the focus of their analysis (e.g., learners inferencing strategies, impact of L1 cultural backgrounds on accuracy, longitudinal gains, etc.).

**Table 1**

*Reviewed L2 indirect meaning studies (5 pages total)*

Author(s) / year	Target Language	Participants	Test Design/Format/Measure	Types of Implicatures (items) or speech acts	Research Focus
Takahashi, Beebe 1987	English	20 Japanese learners of English in Japan 20 Japanese learners of English in the US 10 Japanese native speakers in Japan 20 English native speakers in the US	Written Discourse Completion Tasks with 12 situations to elicit refusals	Indirect refusals (for each group the refusals were addressed to a person of lower, equal or higher status) to: - requests (3) - invitations (3) - offers (3) - suggestions (3)	Analysis in the frequency of typical formulas used for the different types of refusals (in different contexts) and the differences between NS and NNS.
Takahashi, Roitblat 1994	English	23 Japanese learners of English 59 native speakers of English	<i>Experiment 1 (pilot)</i> : Written Test with 26 stories: 12 practice fillers and 14 target stories. Each story was paired with two options paraphrasing either the literal meaning or the indirect request of the final sentence. Followed by a 5-point scale to rate the effectiveness of the story context to interpret the meaning (literal or request) <i>Experiment 2</i> : Written test with 15 stories: 4 practice, 8 fillers, 6 experimental. Each story contained 5 context-inducing sentences (either request-inducing or literal-inducing context) followed by one ambiguous sentence (prime sentence) which participants were then asked to interpret as either literal or request (two options)	Indirect requests vs literal interpretation	Effectiveness of contextual information in interpretation of literal utterances or indirect requests.
Koike 1996	Spanish	46 beginners: Spanish learners in their second semester of intensive Spanish program 34 intermediate: Spanish learners at the end of their third semester of intensive Spanish program 34 advanced learners	Audio-visual test with 7 situations Followed by three questions: 1) use own words to say what the speaker said 2) MC question with list of speech acts (invitation, apology, Request, Suggestion, Order, Information question, Mild Rebuke, Other (specify)) followed by a request for the participant to write how the speaker said the last sentence 3) 5-pt scale rating the speaker's attitude/intent (aggressive/passive, rude/polite, unfriendly/friendly, etc.)	indirect requests (6) literal interpretation (6)	Impact of context (literal-inducing or conventional request-inducing) on processing (reading) speed of the contextual sentences as well as the prime sentences. Differences between NS and NNS
Maeshiba, Yoshimaga, Kasper, Ross 1996	English	30 Japanese learners of English (Intermediate) 30 Japanese learners of English (Advanced) 30 Japanese native speakers 100 learners enrolled in an ESL program at an Australian University: - 50 high proficiency = 25 Japanese and 25 Chinese native speakers - 50 low proficiency = 25 Japanese and 25 Chinese native speakers 50 native English speakers	Written Dialogue Construction Questionnaire with 20 dialogues to elicit apologies (both the offender's and offended party's turns).	Speech acts: Request (1) Information question (1) Rebuke (1) Suggestions (4)	Analysis in the frequency of typical formulas used for the different types of refusals (in different contexts) and the differences between NS and NNS.
Cook, Liddicoat 2002	English	100 learners enrolled in an ESL program at an Australian University: - 50 high proficiency = 25 Japanese and 25 Chinese native speakers - 50 low proficiency = 25 Japanese and 25 Chinese native speakers 50 native English speakers	Written Test with 15 items Multiple-choice questionnaire with 4 options	Apologies in different social contexts and relationship status based on gender, social status, etc., as well as differing degrees of offense severity.  Direct Request (literal) (5) Conventional Indirect Request (5) Nonconventional Indirect Request (5)	Impact of contextual factors Analysis of potential negative or positive L1 transfer on evaluation of contextual factors and L2 production of apology strategies.  Impact of proficiency on comprehension of different request types (literal, non-literal conventional or not conventional) Impact of directness and conventionality on processing and strategies.

Author(s) / year	Target Language	Participants	Test Design/Format/Measure	Types of Implicatures (items) or speech acts	Research Focus
Yamanaka 2003	English	43 ESL learners divided into 4 different levels of proficiency and three different length of residence 13 native speakers of American English	Proficiency test (cloze test) Audio-visual test with 14 items: 2 practice and 12 experimental vignettes Multiple-choice questionnaire with 4 options Open-ended question about contextual clues	Conversational implicatures: Facetious/sarcastic question (1), sarcasm (2), implied negative evaluation (1), irony (1), cynicism/irony (1), rhetorical question (1), violation of relevance maxim (2), parody (1), implied negative comparison (2)	Differences in NS and NNS comprehension of implicatures Impact of proficiency and length of residence in the US on the comprehension of different implicatures Analysis of learners' strategies/use of contextual clues
Garcia 2004	English	56 total: 16 non-native high-proficiency learners 19 non-native low-proficiency learners 21 native speakers	Listening Test with 79 items with only 12 items being analyzed and reported on. Multiple choice questionnaire with 4 options.	Nonconventional indirect meaning: Indirect requests Indirect suggestions Indirect corrections Indirect offers	Impact of proficiency on the identification of different speech acts (including comparison with NS performance). Analysis of potential factors that may have hindered or facilitated comprehension
Shively, Menke, Manzon-Omundson 2008	Spanish	55 L2 Spanish learners enrolled in an American university: - 20 in their second semester - 17 in their fourth semester - 18 in their sixth semester of Spanish instruction	Written test and one video-enhanced test with 8 scenes based on movies (5 target and 3 distractors). Three questions: 1) What does X mean by this comment? 2) A MC question focusing on the "tone" of the last comment (encouraging, authoritative, sincere, sarcastic, sad, etc.) 3) Whether the participant has seen the movie before.	Irony utterances (5)	Impact of L2 proficiency on irony interpretation Impact of the presence of contextual cues (audio-visual format) on L2 comprehension of irony (comparison of written-only instrument with the audio-visual version)
Shively 2013	Spanish	1 American student participating in a study abroad program in Spain	Audio recordings of naturalistic conversations in Spanish by the student Analysis of transcript according to different types of humor, functions, mechanisms, contextualized cues, topics and presence/absence of revoicing)	Humor utterances (5 genres were coded in the transcript: joking about oneself, about absent other, or object, teasing and word play).	Changes in use of humor by a L2 learner during the course of a study abroad Analysis of failed humor Analysis of the ways the student were socialized into their L2 humor practices
Kim 2013	English	28 Korean adults with advanced level English proficiency	Video clips from <i>Friends</i> Written test with 3 tasks (sarcasm identification task, speaker intent comprehension task, potential sarcasm cue identification task) Follow-up verbal interviews with each participant	Sarcasm	Investigate EFL speakers (Korean native)'s perception and detection of sarcasm in English. Identification of detection strategies and potential difficulties in order to present a pedagogical approach to teach sarcasm awareness and interpretation
Bouton 1988	English	436 non-native speakers enrolled in an American university: - 30 German - 32 Spanish/Portuguese - 73 Chinese Taiwanese - 71 Korean - 28 Japanese - 68 Chinese Mainland 28 native speakers	Written implicateure test with 33 items Multiple-choice questionnaire with 4 options	Conversational implicatures based on various maxims among which (two items were missing from the Bouton's list): Relevance (13), Pope Q (2), Quality (1), Change subject (Relevance) (1), Irony (Quality) (4), Manner: Sequence (2), Quantity (4), Negative Evaluation (Quantity) (4)	Impact of L2 learners' L1 cultural background on comprehension accuracy across all implicature types. Differences in comprehension between NS and NNS across different implicature types. Identification of patterns in selected incorrect options on MCQ

Author(s) / year	Target Language	Participants	Test Design/Format/Measure	Types of Implicatures (items) or speech acts	Research Focus
Bouton 1992	English	30 non-native speakers enrolled in an American University (part of the group used for the 1988 paper) 28 native speakers (data from Bouton, 1988)	<i>Longitudinal Study</i> <i>Pre-test (administered in 1986--data from Bouton, 1988)</i> <i>/ post-test (administered in 1991, 4.5 years after arrival)</i> Proficiency test tests including: - Cloze test - Structure test - Dictation test Written Implicature test with 33 items (same test as Bouton, 1988) Multiple-choice questionnaire with 4 options	Same conversational implicatures as Bouton (1998) - one Scalar Implicature was added - three fillers/distractors (non-implicature) were added	Differences between NS and NNS comprehension across implicature types - Analysis of gains between 1986 and 1991
Bouton 1994a	English	Immersion groups: - 4.5 year group from Bouton (1992) - 17-month group: 34 non-native speakers enrolled in an American University 77 native speakers (control group assessed in 1990)	<i>Follow-up longitudinal study</i> <i>Pre-test Post-test 1990-1992: 17 months later</i> Written Implicature test with 25 items (modified from Bouton, 1992) Multiple-choice questionnaire with 4 options	Modified from Bouton (1998): items with questionable validity were removed or edited: - two Pope 2 were added - one Scalar Implicature was added - three fillers/distractors (non-implicature) were added	Differences between NS and NNS comprehension across implicature types - Analysis of gains between 1990 and 1992 (17 months) - Comparison of gains between the 17-month group and the 4.5-year group
Bouton 1994b	English	Immersion groups (students enrolled at an American University): - 17-month group (from Bouton, 1994a) - 33-month group: Non-native speakers (unspecified number) - 4-7 year group: Group of Chinese students 77 native speakers (control group assessed in 1990)	<i>Comparative follow-up longitudinal study expanding on Bouton (1994a)</i> <i>Pre-test Post-test 1990-1992: 33 months later</i> <i>The test was administered to the 4-7 group as a benchmark</i> Written Implicature test with 25 items (modified from Bouton, 1992) Multiple-choice questionnaire with 4 options	Same conversational implicatures as Bouton (1994a) among which (some were missing from the lists presented in the paper): Relevance (9), Minimum Requirement Rule (2), Indirect Criticism (4), Scalar (1), Pope Q (3), Sequence (1), Irony (3)	Differences in comprehension between NS and NNS across implicature types. Differences in gains between the 17-month and the 33-month groups and comparison with the 4-7 year group across implicature types.
Bouton 1999	English	<b>Experimental group:</b> 14 non-native students enrolled in an Academic English course at an American University <b>control group:</b> two other sections of the same course  <b>experimental group:</b> 55 international graduate students enrolled in an ESL program <b>control group:</b> 109 students enrolled in the other sections of the same course.	<i>Pilot Treatment Study (6 hours over 6 weeks of formal introduction of implicature types with examples, and discussion on formulas, and comprehension/production activities)</i> <i>Pre and post tests were administered to both experimental and control groups</i> Written Implicature test with 22 items (modified from Bouton, 1992) Multiple-choice questionnaire with 4 options  <i>Similar to Bouton 1994b Pilot Treatment Study (6 hours over 6 weeks of formal introduction of implicature types with examples, and discussion on formulas, and comprehension/production activities)</i> <i>Pre and post tests were administered to both experimental and control groups</i> Written Implicature test with 22 items (modified from Bouton, 1992) Multiple-choice questionnaire with 4 options	22 Conversational Implicatures - Relevance (9) - Minimum Requirement Rule (2) - Indirect Criticism (4) - Scalar (1) - Pope Q (3) - Sequence (1) - Irony (2)  22 Conversational Implicatures - Relevance (9) - Minimum Requirement Rule (2) - Indirect Criticism (4) - Scalar (1) - Pope Q (3) - Sequence (1) - Irony (2)	Impact of classroom instruction on implicatures (comparison of experimental and control group). Comparison of immersion groups, experimental groups and NS performance across implicature types.  Impact of classroom instruction on implicatures (comparison of experimental and control group). Comparison of immersion groups, experimental groups and NS performance across implicature types.



Author(s) / year	Target Language	Participants	Test Design/Format/Measure	Types of Implicatures (items) or speech acts	Research Focus
Taguchi 2002	English	8 Japanese learners of English enrolled in an American University = 4 lower proficiency + 4 higher proficiency	Listening test with 24 items: 2 practice, 15 experimental, 7 control (literal) Yes/No/I don't know questionnaire Verbal reports	<b>Indirect replies:</b> indirect refusals (5) indirect opinions (5) disclosures (5)	Impact of proficiency on comprehension accuracy Identification of inferential strategies Impact of proficiency on strategy use across implicature types
Taguchi 2005	English	210 total: 46 native speakers (students at an American University) 160 Japanese learners of English in Japan	listening test with 40 items: 2 practice, 32 experimental (divided into 2 categories), 6 fillers (literal) Multiple-choice questionnaire with 4 options	<b>More Conventional Implicatures (MCIs)</b> = indirect requests (8) and indirect refusals (8) <b>Less Conventional Implicatures (LCIs)</b> = Negative indirect opinions (8) and positive indirect opinions (8)	Differences in accuracy and speed across implicature types (conventional or not conventional). Impact of proficiency on accuracy and speed across implicature types Relationship between L2 accuracy and speed Relationship between L1 and L2 comprehension speed Analysis of MCQ error data
Taguchi 2007	English	112 total: 20 native speakers 92 EFL learners (Japanese native speakers studying English in Japan)	<i>Longitudinal Study</i> <i>Pre-test Post-test (7 weeks of content-based instruction)</i> Lexical Access Test TOEFL Listening test with 60 items: 2 practice, 10 filler, 48 experimental Yes/no questionnaire	<b>More conventional:</b> Indirect refusals (24) = refusal to requests (6), refusals to invitations (6) refusals to suggestions (6), refusals to offers (6) <b>Less conventional:</b> Indirect opinions (24) = negative opinions (12), positive opinions (12)	Development of accuracy and speed across implicature types (conventional or not conventional) Relationship between L2 proficiency, lexical access and implicature comprehension accuracy and speed
Taguchi 2008a	Japanese	84 learners of Japanese enrolled in an American University: - 30 Elementary - 33 Intermediate 22 native speakers (for design)	Listening test with 50 items: 2 practice, 12 filler and 36 experimental Multiple-choice questionnaire with 4 options English Sentence Reading Task Verbal reports with 5 participants each of the two proficiency groups	Indirect refusals (conventional) (12) Conventional indirect opinions (12) Non-conventional indirect opinions (12)	Impact of proficiency on comprehension accuracy and speed across implicature types (conventional or not conventional). Comparison of basic processing speed and pragmatic processing speed. Analysis of MCQ error data Impact of proficiency on inferential process Difficulties in comprehension
2008b	English	44 ESL learners (Japanese native speakers studying English in the US)	<i>Longitudinal Study</i> <i>Pre-test (3 weeks) Post-test1 (8 weeks) Post-test2 (19 weeks)</i> Listening test with 60 items: 2 practice, 10 filler, 48 experimental Yes/no questionnaire Lexical Access Test Language Contact Survey	<b>More conventional:</b> Indirect refusals (24) = refusal to requests (6), refusals to invitations (6) refusals to suggestions (6), refusals to offers (6) <b>Less conventional:</b> Indirect opinions (24) = negative opinions (12), positive opinions (12)	Development of accuracy and speed across implicature types (conventional or not conventional) in a study abroad environment Relationship between L2 proficiency, lexical access and implicature comprehension accuracy and speed

Author(s) / year	Target Language	Participants	Test Design/Format/Measure	Types of Implicatures (items) or speech acts	Research Focus
Taguchi 2008c	English	60 EFL learners (Japanese native speakers studying English in Japan) 57 ESL learners (Japanese native speakers studying English in an American college)	<i>Longitudinal Study</i> <i>Pre-test Post-test</i> Listening test with 60 items: 2 practice, 10 filler, 48 experimental Yes/no questionnaire	<b>More conventional:</b> Indirect refusals (24) = refusal to requests (6), refusals to invitations (6) refusals to suggestions (6), refusals to offers (6) <b>Less conventional:</b> Indirect opinions (24) = negative opinions (12), positive opinions (12)	Differences in development of accuracy and speed between ESL and EFL environments across implicature types (conventional or not conventional).
Taguchi 2009a	Japanese	84 learners of Japanese enrolled in an American University: - 30 Elementary - 33 Intermediate - 21 Advanced 22 <i>native speakers (for design)</i>	Listening test with 48 items: 12 filler and 36 experimental Multiple-choice questionnaire with 4 options Verbal reports with 5 participants for each proficiency level	Indirect refusals (conventional) (12) Conventional indirect opinions (12) Non-conventional indirect opinions (12)	Impact of proficiency on comprehension accuracy across implicature types (conventional or not conventional). Impact of proficiency on inferential process Difficulties in comprehension
Taguchi 2009b	English	48 EFL learners (Japanese students enrolled in Japanese college) 25 <i>native speakers</i>	<i>Longitudinal Study</i> <i>Pre-test Post-test</i> Listening test with 42 items (based on corpora dialogues): 2 practice, 32 experimental, 8 fillers (literal) Multiple-choice questionnaire with 4 options	<b>More conventional implicatures</b> (16) = Indirect refusals (8) and routines (8) <b>Less conventional implicatures:</b> non-literal comments/indirect opinions (16)	Development of accuracy and speed across implicature types (conventional or not conventional).
Taguchi 2011	English	64 Japanese learners of English divided into three groups: - group 1: 22 low proficiency/no study abroad experience - group 2: 20 high proficiency/no study abroad experience - group 3: 22 high proficiency/study abroad experience 25 <i>native speakers</i>	<i>Longitudinal Study</i> <i>Pre-test Post-test</i> Listening test with 42 items (based on corpora dialogues): 2 practice, 32 experimental, 8 fillers (literal) Multiple-choice questionnaire with 4 options	<b>More conventional implicatures</b> (16) = Indirect refusals (8) and routines (8) <b>Less conventional implicatures:</b> non-literal comments/indirect opinions (16)	Impact of proficiency and study abroad experience on the development of accuracy and speed across implicature types (conventional or not conventional).
Taguchi, Li, Liu 2013	Chinese	71 L2 Chinese learners enrolled in a private university in the US divided in three groups: - 21 elementary non-heritage - 25 advanced non-heritage - 25 advanced heritage 22 <i>Chinese native speakers</i>	Listening test with 46 items: 2 practice, 8 filler and 36 experimental items English Sentence Reading Task	Indirect refusals (conventional) (12) Conventional indirect opinions (12) Non-conventional indirect opinions (12)	Differences in comprehension accuracy and speed between implicature types (conventional or not conventional). Impact of proficiency and learning context (FL vs. heritage)
Taguchi, Gomez- Lach, Arrufat- Marques 2016	Spanish	32 L2 Spanish learners (enrolled at a private US college) 21 <i>native speakers</i>	Audio-visual test with 48 items: 2 practice, 36 experimental and 6 literal (fillers). Retrospective verbal interviews with 4 participants from the higher-scoring group and 4 participants from the lower-scoring group.	Indirect refusals (conventional) (12) Irony (conventional) (12) Indirect opinions (nonconventional) (12)	Differences in comprehension between implicature types (conventional or not conventional) Comparison between NS and NNS Comprehension accuracy and speed Identification of inferential strategies Differences in selection and use of strategies between high-scoring and low-scoring groups and based on implicature types

These studies' findings will be discussed in the next three sub-sections as they relate to proficiency and amount of instruction/exposure ([section 3.1](#)), the different types of indirect meaning ([section 3.2](#)), and the inferencing strategies used to interpret the indirect meaning ([section 3.3](#)).

### 3.1. Impact of proficiency on L2 comprehension

As stated previously, we expect proficiency to impact indirect meaning comprehension in the same way it impacts general comprehension in L2. Many studies have demonstrated that increase in proficiency correlates with increased pragmatic competence (in both production and comprehension—cf. Bardovi-Harlig & Dornyei, 1998; Bardovi-Harlig, 1999, 2009, 2012, 2013, 2014; Bardovi-Harlig & Bastos, 2011; Ishara & Cohen, 2010; Kondo, 2010, Levorato, 1993; Rose, 2005; Taguchi, 2005, 2019). The general consensus is that learners must be able to understand the linguistic (literal) meaning of a sentence to reach its correct pragmatic interpretation. Drawing on several studies, Taguchi and Yamaguchi (2019, p. 32) state that:

To infer non-literal meaning, learners draw on their linguistic resources (e.g., grammar and vocabulary), as well as general skills of reading and listening. Unless there are other salient cues that assist comprehension, understanding utterance level meaning is prerequisite to implicature comprehension.

This of course is not without reminding us of the debate over the necessary processing of the literal meaning which we discussed previously (in our section on irony—[section 2.6](#)). Yet, it seems common sense to assume that, for the most part, if learners are not able to understand a sentence at the basic semantic level, they may really struggle to understand its illocutionary force. Therefore, a certain level of linguistic (lexical and grammatical) competence, a “threshold of L2 knowledge and abilities” (Taguchi & Yamaguchi, 2019, p. 32) seems to be a prerequisite to understand indirect meaning (also cf. Rose & Kasper, 2001). What is true for comprehension mechanisms also seems to apply for production. Indeed, Kuriscak (2010) states that “higher proficiency learners tend to be more indirect than their lower-proficiency counterparts (who tend to be more direct in their strategies), they use more external modifiers and more words, more downgraders, and may be less influenced by L1 transfer” (p. 32).

In this section, we will review studies that aim to identify the impact of proficiency on the development of the pragmatic competence, and of course, more specifically the comprehension

of indirect meaning. Because different authors may measure proficiency differently (through a variety of proficiency tests, as well as based on academic levels), we will also include studies that focus on amount of instruction or amount of exposure since both usually impact proficiency (and in many cases, the assumption is that proficiency has de facto increased after a certain amount of classroom instruction or exposure).

Among the cross-sectional studies we can reference as evidence of the positive impact of proficiency on the use or ability to understand indirect meaning are:

- *Koike (1996)* found that proficiency affects learners' recognition of the speaker's intention when using speech acts (invitation, apology, request, suggestions, order, etc.). Less proficient learners are less likely to recognize the illocutionary force of the speech acts.
- *Maeshiba, Yoshinaga, Kasper, and Ross (2006)* demonstrated the impact of proficiency on pragmatic transfer. Advanced learners were better at identifying situations in which L1 apology strategies could be used in L2.
- *Takahashi and Beebe (1987)* found that with increased proficiency and ease of communication in the L2, advanced learners were more likely to be subject to negative transfer. Because they had the linguistic skills to express what they wanted to say (as opposed to less proficient learners), they transferred some of their L1 conventional expressions onto L2 (and due to the differences in cultural conventional expressions, were unsuccessful in reaching native-like communication).
- *Cook and Liddicoat (2002)* looked into L2 comprehension of request intentions and found that as proficiency develops, comprehension of conventional indirect expressions increases.
- *Shively, Menke, and Manzon-Omundson (2008)* focused on irony comprehension by L2 Spanish learners, using video clips. They found that high-proficiency learners were usually able to detect the ironic intent in the speaker's utterance, but lower proficiency learners tended to take the comments literally and failed to recognize the discrepancy between what was said and what was implied (they did not really look for contextual cues and devoted most of their attention to the sentence-level meaning of the utterance).

- *Shively (2013)*: In her study on L2 humor<sup>13</sup> development, Shively (2013) found that over time, L2 learners who participated in a study abroad program showed a decrease in failed humor.

- *Paris (2015)*: in her doctoral dissertation, Paris hypothesized that lower-proficiency learners are less likely to use non-literal language as their working memory and attention efforts are focused on micro-programming. Both the exploratory and comparative studies she conducted showed that higher-proficiency learners were able to correctly use more native-like non-literal expressions as compared to lower-proficiency learners.

Of course, studies that compare learners' production or comprehension of indirect meaning to native speakers' performance may also provide evidence that L2 proficiency plays a big role in understanding indirectness. For example, Takahashi and Roitblat (1994) showed that L2 learners took longer to comprehend indirect requests when compared to native speakers. Taguchi (2009b) found that native speakers' comprehension of implicatures was almost perfectly accurate and much faster than EFL learners (even after over six years of formal EFL instruction).

Yet, a higher level of proficiency is not always an indicator of success when it comes to indirect meaning comprehension. While we may assume that a certain level of grammatical and lexical knowledge is necessary to comprehend a non-literal utterance, Bardovi-Harlig (1999) explains that "high levels of grammatical competence do not guarantee concomitant high levels of pragmatic competence" (p. 686). Along the same lines, Ishara and Cohen (2010) state that (p. 80):

Learners who can understand and produce highly accurate language forms from a grammatical point of view are not necessarily able to use language in a pragmatically appropriate manner. [...] Conversely, learners who demonstrate very little grammar accuracy may still be able

---

<sup>13</sup> Shively (2013) defines humor comprehension in much the same way some of the authors we discussed in part II of this chapter define irony and irony processing. Specifically, she states that (2013, p. 931): "Humor is generally believed to be derived from incongruity or script opposition; that is, the juxtaposition of elements in a way that is perceived to be contradictory or inconsistent with expectations (e.g., Attardo, 2001, 2008; Raskin, 1985). In this view, script refers to the cultural expectations, roles, and settings associated with particular words and phrases. To understand the incongruity underlying humor, a learner must be able to access the scripts that are evoked, which may require broad knowledge about the target culture (Chiaro, 1992, Yus, 2004). The learner also needs to understand the logical mechanism causing the script opposition, be familiar with situation and target, and understand the narrative strategy, as well as the language that forms the exact wording of the humor (see Attardo, 2001)."

to interpret messages as intended and produce pragmatically appropriate utterances.

On the other end of the spectrum, as suggests the second half of Ishara and Cohen's (2010) quote above, lower proficiency level does not necessarily mean that learners cannot successfully comprehend the illocutionary force of an utterance and understand the intended message or even attempt to use an expression before they have acquired it.

We will now turn our attention to a few key studies which specifically look into the three different types of meaning we are focusing on for this dissertation: indirect refusals, indirect opinions and irony. In particular, we will focus on whether these authors found a correlation between proficiency and comprehension of these types of implicatures.

**Bouton (1992):** this longitudinal study investigates the link between L2 proficiency and implicature comprehension as well as the impact of exposure to the target language (participants were tested in 1986 and then again in 1991). Proficiency was measured using a combination of three tests: a structure test, a cloze test and a dictation test. Implicature comprehension was assessed with a multiple-choice questionnaire. Participants were presented with written dialogues, which he based off of another study that he published in 1988 (which we will discuss later), in which he included implicatures based on the Relevance Maxim (including Pope Question), implicatures based on the Quantity Maxim, such understated criticism (which he called understated negative evaluation in his 1988 paper—and which we assimilate to the negative indirect opinions in our dissertation), and implicatures based on the Quality Maxim (i.e., irony).

As expected, overall proficiency was higher in 1991 than in 1986 (likely the result of the participants living in a target language environment during those five years). Bouton finds that there was only a weak relationship between the proficiency test results and implicature comprehension, both in 1986 and 1991. He concludes that “We cannot measure a person's ability to interpret implicature by using a general proficiency test like the EPT” (p. 56). However, he does note some gains between the 1986 and 1991 implicature comprehension tests. While the results of the non-native speakers were still significantly lower than those of native speakers, their comprehension had significantly improved between 1986 and 1991 and after 4.5 years in the country, they were able to correctly interpret most of the implicatures (from understanding 79.5% of the time as compared to native speakers in 1986 to understand 92% of native speakers' correctly interpreted implicatures in 1991). Yet, differences remained as to which types of

implicatures showed the most—or little—growth. We will discuss the difference in types of implicatures in the next section.

**Bouton (1994a):** Following up on the study he presented in his 1992 paper (discussed above), Bouton investigates non-native speakers' comprehension of conversational implicatures after 17 months in the country, and compare these results to the findings he reached in the 1992 paper with a group of non-native speakers who had been in the country for about 4.5 years. He acknowledges some changes made to the implicature comprehension test (the test was reduced from 33 items to 25, some items from the original test were eliminated due to their questionable validity, and two Pope Q implicature as well as one scalar implicature were added. He also included three non-implicature items as distractors). Results show an increase in implicature comprehension after 17 months in the country. However, the non-native speakers' scores were still significantly lower than those of the native speakers, even after 17 months. More interestingly, it was clear that some types of implicatures were still considerably challenging for non-native speakers who had only been in the country for 17 months (whereas, by 4.5 years, they had fewer difficulties<sup>14</sup>). He concludes that while living in the US for an extended period of time does improve comprehension of implicature, this is a long process without guided learning “in this area of cross-cultural communication” (p. 167).

**Bouton (1994b):** this multi-level paper expands on the two previous studies. This purpose of this study was to establish a timeline or find out the amount of time or exposure that is required to develop a certain level of proficiency in implicature comprehension. Thus, the researcher tests non-native speakers upon their arrival in the US, after 17 months in the country, and then another group after 33 months. Bouton also includes results from a group of EFL learners who had been in the US for 4 to 7 years (he explains that he could not use the results from the 4.5 years group in the previous studies due to the instrument being different). Statistically significant differences were found between arrival and 17 months and arrival and 33 months but no statistical differences were observed between 17 and 33 months (or with the 4-7 years group<sup>15</sup>). All non-native speaker groups differed from the native speaker group. Bouton concludes that most of the progress in implicature comprehension was achieved during the first 17 months and not much improvement was observed after that (the difference between the 4-7 years

---

<sup>14</sup> However, we know that the implicature comprehension test used with the 17-month group was different than the one used in his first study (with the 4.5-year group).

<sup>15</sup> It is interesting to note here that this seems to go against Bouton's previous findings (1992) since he had observed differences between the 17-month and the 4.5-year groups.

group and native speakers were still significant). Yet again, some types of implicatures were resistant to comprehension even after 7 years in the country, but “the longer the international students were in the United States, the fewer the number of implicature types that were systematically problematic for them” (Bouton, 1999, p. 59). As we will discuss in the next section, the degree of conventionality or opacity of the implicatures seemed to factor into the potential for developing comprehension. The pilot treatment study in Bouton (1994b) also showed that some types of implicatures responded to instruction better than others (which we will discuss later).

**Taguchi (2002):** Drawing on the Relevance Theory as her framework, Taguchi investigates EFL learners’ inferential ability and the impact of proficiency on comprehension. She found that, while her higher proficiency group performed better overall on the comprehension of indirect meaning (indirect opinions, indirect refusals and disclosures) than the lower-proficiency group, both groups were able to understand implicatures and the difference in scores seemed to have more to do with confidence level than actual comprehension. Referencing strategies that we will discuss later, she states (2002, pp. 160-161):

Inference processing also seems to function similarly regardless of L2 proficiency. Learners with limited proficiency can still utilize context and search for the most relevant interpretations of indirect utterances. These findings imply that inferential communication and the search for relevance are part of ordinary cognitive ability and can be transferred to L2 comprehension process.

Taguchi also observed that there was no significant difference between the low and high proficiency groups in terms of the number of strategies they used to interpret the implicature. However, proficiency did impact the type of strategies learners used. We will discuss inferencing strategies in section [3.3](#), but Taguchi’s results revealed that low-proficiency learners used keyword inferencing as the most common processing tool, and at a much higher rate than the high-proficiency learners. Kasper (1984) and Ross (1997) had already shown that until learners reach a higher level of proficiency, they may rely on bottom-up processes based on syntactic and lexical information at the expense of more complex pragmatic inferencing processes. The high-proficiency learners in Taguchi’s (2002) study however used more strategies based on contextual cues (paralinguistic, adjacency pair rule, and speaker intention).



**Yamanaka (2003):** This study is one of the few<sup>16</sup> that uses audio-visual input as part of their instrument. Yamanaka used excerpts from television series to create an instrument to assess EFL learners' comprehension of different types of implicatures, such as irony, sarcasm, relevance-based violation, negative evaluation and comparison, rhetorical questions and parody. Yamanaka focused on two central ideas: first she wanted to determine the impact of proficiency (measured by a cloze test, and dividing her non-native speakers into four different levels as well as comparing with native speakers' performance) and length of residence (and which of the two factors—proficiency and residence length—was a better predictor of success with indirect meaning comprehension). Second, she looked into the learners' inferencing process: she asked participants to note the type of contextual clues they attended to in order to find out whether failed comprehension was due to inability to notice certain clues (based on reported clues from native speakers).

Not surprisingly, the difference between native speakers and non-native speakers was overall statistically significant. Yamanaka also found a strong correlation between proficiency and accuracy as well as a positive impact of length of residence (though the correlation coefficient for proficiency was higher than that for length of stay). Bouton (1992) had previously shown a weak relationship between proficiency and pragmatic comprehension, which led Yamanaka to divide her non-native speaker groups into four proficiency levels<sup>17</sup> based on their cloze test scores. Again, all non-native speaker groups had significantly lower scores than the native speaker group. However, within the learner groups, differences were only found between group 1 (16 learners who had scored less than 25% on the cloze test) and 4 (nine learners who scored between 75 and 100%) and between group 2 (nine learners who scores between 25 and 49%) and 4.

Following Bouton's (1994a) study comparing non-native speakers' performance at 17 months and 4.5 years in the US, Yamanaka divided her non-native speakers into three

---

<sup>16</sup> In addition to Shively, Menke, & Manzon-Omundson (2008), Taguchi et al. (2016), Kim (2016), Kim & Lantolf (2016) who used video clips for their studies, other studies were either based on written test (c.f. Bouton's studies) or listening tests (cf. most of Taguchi's studies we reviewed).

<sup>17</sup> Proficiency testing was done differently in Bouton's study (he used TOEFL scores as well as a battery of tests, including cloze test) and Yamanaka (only used the cloze test). Yamanaka believed Bouton's participants had a higher level of proficiency overall (the TOEFL scores he reported were fairly high) which may have impacted his results. Based on Bouton's data, we know that his cloze test did not use the same number of items (he reports a mean of 56 and Yamanaka indicated that her cloze test included 50 items), but he did find a very small correlation between cloze test results and implicature comprehension in 1986 (which had actually weakened by 1991). In her paper, Yamanaka points out some of the limitations and contradictions of Bouton's studies when it comes to the effect of proficiency.

groups based on these lengths of residence: short-term (0 to 17 months), medium-term (18-53 months) and long-term (over 54 months). Results show a significance difference in performance for the short and medium-term group as compared to the native-speaker group, as well as between the short-term group and the long-term group. Yet, the shared variance between proficiency (cloze test) and non-native speaker scores on the implicature questionnaire was higher than that of the relationship between scores and length of residence. In the discussion section of her paper, Yamanaka provides some insights into the comprehension process of different types of implicature, which we will discuss in our next section.

**Taguchi (2005):** Taguchi investigates the impact of ESL learners' proficiency on the comprehension of implied meaning using implicatures of different degrees of conventionality (more conventional implicature: indirect refusals and indirect requests; and less conventional implicature: negative, positive or neutral indirect opinions). The instrument includes a listening test followed by a multiple-choice questionnaire with 4 options: one correct answer with the intended meaning, and three distractors written according to the following principles (p.550):

Principle 1: The option contains a meaning that is the opposite of the implied meaning.

Principle 2: The option contains words taken from the last part of the dialogue.

Principle 3: The option is related to the overall conversation.

Results show that L2 proficiency<sup>18</sup> impacts accuracy scores for all types of conventional implicature but differences in comprehension speed were also observed for more conventional implicature. However, the size of the coefficient was small for response time which led Taguchi to conclude that there is a weak relationship between L2 proficiency and comprehension time. She analyzed the incorrect answers to determine which distractors attracted the low-proficiency and high-proficiency groups the most. She found that overall, regardless of their proficiency, Principle 2 was the option that L2 learners seemed to choose more frequently when selecting the wrong answer. Taguchi made the case that this was due to the adjacent pair effect and learners' working memory: "when implied meaning is less accessible to learners, they tend to rely on their memory more in order to make choices" (p. 555). However, we argue that this may also draw on keyword inferencing strategy, especially for lower-proficiency learners who may focus

---

<sup>18</sup> L2 proficiency was measured using institutional TOEFL scores.

on specific words in the dialogue due to their inability to fully understand the conversational exchange (which could impact comprehension in general). One way to determine whether comprehension difficulties are a result of indirect meaning only (as opposed to proficiency impacting overall comprehension but not specifically indirect meaning) would be to test comprehension of literal meaning items using the same instrument. Taguchi did use six literal items as fillers for her study, but she did not analyze the results or include a comparison of comprehension of these items in her study.

**Taguchi (2007):** In this study, Taguchi looks into ESL learners' comprehension of indirect refusals and indirect opinions in terms of accuracy and speed, over a period of seven weeks of intensive ESL instruction (which did not focus on implicature comprehension). She also examines the link between L2 proficiency (using the ITP TOEFL), speed of lexical judgment, and pragmatic comprehension ability. She found that L2 learners' comprehension had improved significantly after seven weeks, though accuracy had shown more significant gains than speed. Her findings also revealed that there was a significant relationship between proficiency and accuracy. However, proficiency did not correlate with comprehension speed. Only lexical access speed correlated with comprehension speed (and there was no significant correlation between accuracy scores and response time). Taguchi concluded that "development of pragmatic knowledge and processing capacity of using the knowledge may not coincide perfectly in L2 development" (p. 313).

**Taguchi (2009a):** Unlike the previous three studies we reviewed, this study looks into learners' comprehension of L2 Japanese. Taguchi focuses on three types of indirect meaning: indirect refusals (which she classifies as conventional due to its predictable pattern: providing an excuse), conventional indirect opinions (using specific Japanese conventional expressions, as well as a Japanese-language-specific questioning strategy to show disagreements) and non-conventional indirect opinions. A discussion on conventionality will be presented in the next section, but Taguchi's research questions focused on the impact of proficiency on accuracy (measured by a multiple-choice questionnaire paired with written dialogues) as well as comprehension processing strategies (observed through follow-up introspective verbal interviews). Participants were grouped into three different proficiency levels (EJ: elementary Japanese students who were in their second semester of Japanese instruction; IJ: Intermediate Japanese students in their fourth semester; and AJ: advanced students in their sixth semester). Results showed significant differences in accuracy across all types of indirect meaning,

though post-hoc analyses revealed that only the lowest-proficiency group (EJ) was significantly lower than the two other proficiency groups (IJ and AJ). Proficiency also impacted comprehension processing. Taguchi's verbal interview data showed that the lower-proficiency learners struggled with their listening ability:<sup>19</sup> "the EJ learners mentioned that they could not understand the words, phrases, or situations in the conversations" (p. 260). Taguchi states that (p. 260-261):

Basic comprehension ability seemed to have played the most prominent role in pragmatic comprehension because the EJ learners did not have any problem with understanding the illocutionary force of indirect refusals when they could comprehend the target utterances.

On the other hand, she noted that the more advanced learners relied more on pragmalinguistic rules and were able to verbalize their reasoning during the retrospective interviews (even though some features were still challenging as even more proficient learners were not always familiar with them). In the discussion part of her paper, she highlights the idea that a certain threshold of linguistic ability must be reached for the inferential process to take place, and that once this threshold is met, differences in L2-pragmatic comprehension are not as noticeable between proficiency groups (between IJ and AJ) due to their ability to transfer inferential processes from L1 to L2: "Once the linguistic part of the challenge is controlled, learners seem to be able to transfer their L1-based inferential skill to L2, and successfully seek relevance of the indirect information" (p. 265).

**Taguchi (2011):** this study examines both the impact of general proficiency as well as study-abroad experience on pragmatic comprehension: Group 1 was the lower proficiency group (first semester EFL students in an intensive immersive program in Japan), group 2 had a higher proficiency and no study-abroad experience and group 3 had high proficiency as well as study abroad experience. Similarly, to her other ESL/EFL studies, a pragmatic listening test<sup>20</sup> was administered to assess the different groups' comprehension of conventional implicatures (indirect meaning conveyed by "fixed

---

<sup>19</sup> Taguchi (2009a) specifically blames learner's listening ability here (since the test used for this study was a listening test). However, we may wonder if this was also partially a semantic ability issue. We know that low proficiency learners may not have a very high lexical knowledge, and given that her EJ group was only in their second semester of Japanese instruction, it would not be surprising if there were words or expressions in the dialogue that they failed to understand, despite the fact that Taguchi used vocabulary and grammar "drawn from the Elementary Japanese course packet used in the target institution" (p. 256).

<sup>20</sup> The same test she developed for her 2009 study (Taguchi, 2009b) based on corpora.

linguistic forms or through predictable discourse patterns,” Taguchi 2009b, p. 741—such as indirect refusals and routines) and nonconventional implicatures (“nonliteral comments or opinions that do not involve conventional linguistic features or language use patterns,” Taguchi, 2009b, p. 741). Statistical analyses showed that proficiency (but not study abroad experience) positively affected comprehension speed (for both conventional and nonconventional implicature) as a significant difference was found between group 1 and group 2. In terms of accuracy, group 1 obtained significantly lower accuracy scores than the other two groups, and group 3 score was higher than group 2 (but all three learner groups’ performance was still lower than the native speaker group) in the comprehension of nonconventional implicatures. Conventional implicature comprehension however seemed to only be challenging for group 1. Therefore, study abroad experience did not seem to impact comprehension of more conventional implicature. Further distinctions were made between the comprehension of various implicature (indirect refusals vs. routines for example), but we will address this in the next section.

As we can see from the results of the studies we reviewed above, factors such as proficiency, type of implicature and conventionality, study abroad experience, and target culture all factor into the comprehension of indirect meaning. As Taguchi states (2011, p. 926), there is an “intricate relationship among item type, comprehension ability, and the factors affecting them. Different degrees of comprehension load encoded in implicatures were affected differently by proficiency.” Several times in the above review, we mentioned that while proficiency clearly impacted learners’ comprehension of indirect meaning, it impacted it differently based on the type of implicatures they were trying to interpret. We also briefly mentioned that the types of strategies learners used to comprehend indirect meaning might be different based on their proficiency. In the next section of our L2 literature review we will focus on the types of implicatures analyzed in different L2 studies and how their comprehension may differ based on the degree of conventionality ([section 3.2](#)). We will then review studies that focused on the use of specific comprehension strategies by L2 learners ([section 3.3](#)).

### 3.2. Impact of indirect meaning types on L2 comprehension

Our dissertation focuses on three different types of indirect meaning: indirect refusals, indirect

opinions and irony. The study we are replicating categorizes these types of indirect meaning based on whether they are conventional (indirect refusals and irony) or not (indirect opinions). While reviewing L2 research literature, we wanted to include, not only experimental studies that focus on our three types of indirect meaning (indirect refusals, indirect opinions and irony—whether the studies label them as such or not), but also look into studies that discuss the concept of conventionality—or related notions—and its impact of on L2 comprehension of indirect meaning.

When it comes to indirect meaning, some expressions, implicatures or speech acts seem easier to comprehend than others. But why is that? Are certain implicatures easier to understand than others? Are there factors or features specific to certain types of indirect meaning that make them more accessible to L2 learners? As we have seen previously, some authors refer to a certain degree of indirectness (Thomas 1995), strength of implicature (e.g., Bouton 1992, 1994a, 1994b; Sperber & Wilson, 1981), or level of conventionality (Cook & Liddicoat, 2002; Taguchi, 2002, 2005, 2007, 2008, 2009a, 2009b, 2011; Yamanaka, 2003) that may hinder or facilitate interpretation of indirect meaning. Weizman (1985) speaks of transparency or opacity of the illocutionary force.

The first thing to note is that not all authors agree on the classification of the different types of indirect meaning. Bouton for example, heavily relied on Grice's Maxims as the basis for the classification of his implicatures. We believe that the three types of indirect meaning (indirect refusals, indirect opinions and irony) on which our dissertation focuses were included in Bouton's studies, though it is not always obvious because he did not provide a complete list of all the implicatures he studied. Bouton conducted a series of now-famous experiments (cf. [Table 1](#)) on L2 implicature comprehension (1988, 1992, 1994a, 1994b, 1999). In Bouton (1988) for instance, implicatures were based on the following Maxims: *Relevance* (including *Pope Question*,<sup>21</sup> and *change of subject*), *Quality* (including *Irony*), *Manner* (*Sequencing*), and *Quantity* (including *understated negative evaluation* and *indirect criticism*, which are similar to our negative indirect opinions). One of the examples of Relevance-based implicature Bouton

---

<sup>21</sup> Pope Q: "a violation of the Relevance Maxim based on the prototype dialogue in which the apparently irrelevant question, "is the Pope Catholic?" is given as a response to another question to which the answer seems to the speaker to be "Obviously!" If the speaker wants to use this device to indicate that the answer is "Obviously not!" he chooses as his apparently relevant question one to which the answer is obvious but "No!," e.g., "Is the sky green?" It may or may not be sarcastic." (Bouton, 1988, p. 191).

(1994b) provides is essentially an expression of indirect refusal.<sup>22</sup> It is necessary to point out that Bouton using “Relevance Maxim” as an overall category seems somewhat problematic to us (and we estimate that about a third of the implicatures he used in his studies were classified as such). As we have explained in [Chapter 2](#), Grice’s Maxims were not perfect, and Sperber and Wilson actually argued that Relevance is at the center of all implicature comprehension process (which means all conversational implicatures used by Bouton could be considered “relevance-based”). To complicate the matter, Bouton uses certain maxims (e.g., Relevance, Quantity, Quality) as general categories of implicature, but also to explain some of the more specific implicature (e.g., irony is labeled as an implicature based on the Maxim of Quality, negative evaluation based on the Maxim of Quantity, etc.). Finally, we noted a lack of consistency in his categorization. In Bouton (1994a), one of his examples of Irony is listed as a Relevance Maxim implicature, when it was classified as a Maxim of Quality implicature in Bouton (1988). Further evidence of this problematic classification of implicatures was brought to light by Yamanaka (2003) who argues that the Pope Question (which appears as its own category in Bouton (1994b, 1999) falls under Gibbs’s (1994) description of sarcasm, which, as we have seen in the previous part ([section 2.6](#)), is considered by many a form of irony. Of course, this overlap or unclear classification makes determining if one category of implicature or indirect meaning is more challenging than another much more complicated.

We will now turn our attention to some of Bouton’s (1988, 1992, 1994a, 1994b, 1999) findings to see if patterns of comprehension difficulties emerge based on implicature types. Overall, Bouton’s studies showed that the implicatures that both native and non-native speakers found easiest were those in the Relevance-based category—that were “more generally based on the tendency of participants in a conversation to assume that whatever a person says is somehow related to what has gone before and to interpret it in that light” (Bouton, 1992, p. 59). It is unclear here if what Bouton refers to in the above quote is related to the adjacency-pair rule or more general conversational patterns. He does not use the term conventional or conventionality in his paper, but we may assume that his description refers to a predictable pattern of discourse, following the Relevance Theory. Among those implicatures that were more challenging for non-native speakers—which Bouton referred to as the “Tough 10” (1999)—were *understated (or indirect) negative evaluation/criticism* (1988, 1992, 1994a, 1994b, 1999), *irony* (1988, 1992, 1994a, 1994b, 1999), *Pope Question* (1994a, 1999), implicatures based on the *Maxim of*

---

<sup>22</sup> There is no way for us to tell whether indirect refusals may have also been part of the Relevance implicatures in Bouton’s other studies.

*Quantity* (1988), *Scalar implicature* (1994b, 1999), and *Sequence of events* (1994b, 1999). Yet in his longitudinal studies (1992, 1994a, 1994b) Bouton observed some comprehension gains for some of the above implicatures that were initially hard for non-native speakers to understand. Indeed, after a certain amount of exposure, living in the target language, non-native speakers' performance on some of the more challenging implicatures was similar to that of native speaker, with the exception of implicatures "related to specific points of American culture" as "the problems caused by these items were arbitrary and idiosyncratic" (1994b, p. 163).

Consistent with Bouton's findings (especially with regard to irony) were those of Yamanaka (2003) and Kim (2013, 2014, 2016). Among the types of inferences Yamanaka used for her study were implied negative evaluations, violation of *maxim of relevance*, *irony*, *parody*, *rhetorical questions* and *sarcasm*. Opting to include her rhetorical question item (Pope Q) and sarcasm items in the irony category (based on Gibbs', 1994, description), she concludes that this category of inferences was overall "among the most difficult" (2003, p. 138-139), but notes that "irony can be learned" (2003, p. 139) as she observed an increase in accuracy among the highest proficiency group and long-term residents. Kim's (2013, 2014, 2016) studies focused on one very specific type of indirect meaning: sarcasm. Drawing from previous research, Kim postulates that "L2 learners will process sarcastic utterances through the lens of their L1 knowledge system or schema" (2014, p. 194). She goes on to explain that schemata are complex cognitive systems that "emerge out of knowledge among members of the same language-cultural (Agar, 2002) group over time" (2014, p. 194). Using video clips from the American series *Friends*, she led a group of advanced-level EFL speakers (Korean natives) through a series of tests to find out whether they were able to detect sarcastic comments and whether they managed to uncover the intended message of the utterances and their illocutionary force. Her findings show that her Korean participants were usually able to detect and identify the sarcastic comments but they rated them as a lot more offensive than the native speaker group who usually equated the sarcasm to "light-hearted jokes" (p. 198). Through an intervention study using explicit instruction of sarcasm, in L2 English, Kim (2016) and Kim and Lantolf (2016) noted significant improvements in the Korean natives' conceptual understanding of English sarcasm. These findings are in line with what Bouton reported regarding gain in irony comprehension.

One valuable contribution from Bouton (1994b, 1999) is the idea that some implicatures are formulaic and while others are more idiosyncratic. Bouton (1994b, 1999) states that Relevance-



based implicatures are not formulaic because “their interpretation is idiosyncratically dependent on the relationship between a particular utterance and a specific context” (1994b, p. 98). Nonetheless, Bouton’s studies showed that Relevance-based implicatures were—for the most part—fairly easy to understand for non-native speakers. On the other hand, Bouton (1994b, 1999) explain that implicatures like the Pope Q ones are based on a pattern or formula and the hearer must recognize this structure to properly reach the intended meaning. Similarly, indirect criticism is not formulaic in structure but follows a more semantic formula as it is often used to answer a request for an opinion. Even after being in the US for an extended period of time, Bouton’s non-native speakers still had trouble understanding these implicatures that were “based on a formula of some sort – structural, semantic, pragmatic or some combination of these” (1994b, p. 99). Bouton hypothesized that once non-native speakers are familiar with that formula, they should be able to interpret formulaic implicatures. This idea became the focus of two treatment studies (1994b, 1999). Explicit instruction, discussion and production activities focusing on the following implicatures were used for this treatment: Pope Q, Indirect criticism, Sequence of Events, Irony and Relevance-based. Results showed that instruction sped up progress in interpreting implicatures, but some were more amenable to instruction than others. The more formulaic implicatures (Pope Q, Indirect Criticism and Sequence of events) responded well and, students showed significant gains on those items. Irony remained difficult but considerable progress was made (improvements that were also observed in Yamanaka, 2003 and Kim, 2013, as previously discussed). Relevance-based items were however resistant to the approach and the non-native speaker group actually did worse on the post-test. Bouton thus concludes that (p. 106):

[T]he lack of any overall system underlying relevance implicatures in general means that they can be neither taught nor learned systematically. Once the students have developed an awareness of the existence of relevance implicatures, we should deal with them individually in the classroom as the need arises rather than as a whole set. Formulaic implicatures, on the other hand, are inherently systematic and can be approached effectively from this perspective.

We should also remember here that many relevance-based implicatures were actually found to be fairly easy to comprehend by non-native speakers (even upon arrival in the US), so overall, as a general category, it seems that relevance-based implicatures, while harder to teach, are easy to understand.

This concept of formulaic implicatures is fairly similar to what Taguchi and other authors refer

to as “more conventional implicatures” (for example in Taguchi, 2005, p. 547). Before we look into the studies that focus on the role of conventionality in L2 indirect meaning comprehension, it is important to define what conventionality is within the scope of our research and this dissertation (especially in relation to indirect refusals, indirect opinions and irony). Bromberek-Dyzman and Ewert (2010, p. 318) define conventionality as being “related to the frequency of usage of a particular figurative expression and its salience in the speakers’ minds, which is relative to a community of usage.” We have already addressed the issue of Grice’s definition of conventional implicatures, but what Taguchi refers to when she uses the term conventionality related to implicatures, is utterances in which “the speaker’s intentions are linguistically coded or embedded within predictable, fixed patterns of discourse” (2005, p. 545). In her studies, more conventional implicatures were based on “specific linguistic forms or language use patterns” (2005, p. 547). Taguchi’s definition offers some parallels with what Bouton referred to when he mentioned a specific “linguistic pattern associated with the implicatures” (1992, p. 12) or his “formulaic” expressions (1994b, 1999). One example of a linguistically conventional or formulaic expression is the use of “*Do you mind...?*” followed by a verb to make indirect requests. English language users can easily and quickly access the intention of the speaker (an indirect request) when they hear this expression (Blum-Kulka, House, & Kasper, 1989). An example of a conventional indirect meaning that is not based on a specific linguistic form but rather on a conventional pattern of conversation would be indirect refusal (Holtgraves, 1999; Taguchi, 2005, 2009, 2011, etc.), due to its conventionalized pattern of response to an invitation or offer (which are usually constrained by the adjacency pair rule—see [section 2.7](#)): when refusing an invitation, it is common for someone to provide a reason as to why they may not be able to accept the invitation (and that way they may not directly reject the invitation –cf. Politeness theories [section 2.4](#)).

Takahashi and Roitblat (1994) investigated L2 comprehension of indirect requests which they classified as conventional due to the speech acts using a conventionalized pattern of conversation (for example, asking “Can you tell me the time?” which should be conventionally interpreted as a request to be told what time it is,<sup>23</sup> as opposed to a literal question about someone’s ability to read and tell the time). In their study, they found that both native and non-native speakers could reach context-relevant interpretation for both literal and conventional situations. Unlike previous studies from other authors (Ervin-Tripp, Strage, Lampert, & Bell,

---

<sup>23</sup> This conventional interpretation is the “‘preferred’ interpretation” according to Cook and Liddicoat (2002, p. 22).

1987; Kasper, 1984), Takahashi and Roitblat's L2 learners "were not more likely than native English speakers to interpret the sentences literally" (1994, p. 495). One possible explanation for this difference is the fact that L2 participants in their study had already reached a fairly high level of proficiency and were familiar with the American conventionalized ways of making indirect requests.

Similar to Takahashi and Roitblat (1994), Cook and Liddicoat (2002) were also interested in indirect requests but specifically investigated the effect of L2 proficiency on the comprehension of three different types of requests (based on Blum-Kulka's, 1989, categories): direct requests, conventional indirect requests and nonconventional indirect requests. They defined conventional indirect requests as ones that use "fixed linguistic conventions established in the speech community" (p. 21) with more than one function. As an example, they say that using "*Could you + VP?*" is a direct question about someone's ability but is conventionalized as an indirect request in English (Blum-Kulka, 1989). They go on to explain that knowledge of linguistic conventions is not enough and one must also understand "pragmatic conventions used by the speech community to map conventional form meaning relationships" (p. 22). On the other hand, the nonconventional indirect request is based on a "partial reference to the requested act, object or recipients' involvement" (p. 22). To understand the illocutionary force of these types of requests, one must "use contextualised conventions such as social and contextual knowledge" (p. 22). In other words, other contextual cues must be used to interpret nonconventional indirect requests and, in their study, Cook and Liddicoat included pre-requests and statement hints in their nonconventional indirect request category. Their results showed an effect of both level of directness and conventionality. Indeed, "as the type of request strategy increased in difficulty and decreased in conventionality, as was the case with nonconventional indirect request scenarios, the number of expected interpretations for the L2 English learners decreased dramatically" (p.27). Cook and Liddicoat's lower-proficiency group was able to comprehend direct requests better than indirect requests. Within the indirect request category, the lower-proficiency participants found conventional indirect requests easier to interpret than nonconventional ones. Cook and Liddicoat highlight the relationship between proficiency (and automaticity—cf. Takahashi, 1990, 2010), conventionality (or types of speech acts), and processing capacities (p. 29):

[L]anguage learners may have less capacity available for processing information because of their greater reliance on controlled processes for lower level language tasks. It seems that the variability of learners' performance can be explained by the difference in the learners' capacity

to access their linguistic and contextual knowledge. Native speakers have sufficient capacity to access both linguistic and contextual knowledge and activate both the low (linguistic rules) and the high order frames (contextual knowledge) [...]. This is because they have achieved high levels of automaticity in the processing of language.

The link between conventionality and processing effort, L2 proficiency, and culture is one that we had addressed in the previous section already ([section 3.1](#)), and something that Taguchi highlights regularly in her studies (2002, 2005, 2007, 2008a, 2008b, 2008c, 2009a, 2009b, 2001—cf. [Appendix A](#)). The more proficient a learner will be, the more familiar they will be with L2 linguistic or cultural conventions:

When learners are aware of the conventions or when conventions are shared between first language (L1) and the L2 (e.g., giving a reason for refusals in Japanese and English), meaning is more easily accessed as long as learners have sufficient linguistic skills to comprehend the utterances.

Taguchi, 2008c, p. 428

Taguchi then continues to explain the potential processing difficulties that arise when learners are faced with nonconventional or unknown forms of implicatures (2008c, p. 428):

[W]hen conventions are not familiar or when implied meaning is conveyed through nonconventionalized expressions, comprehension load increases and extensive inferential processing is required, as learners use syntactic and lexical information as well as contextual cues and knowledge of conventions to infer meaning.

Echoing this relationship between proficiency and conventionality, Takahashi (1990) discusses the notion of automaticity. Indeed, we assume that for proficient language users, conventional expressions or routines lead to an automatized processing of the utterance that does not necessarily involve processing the literal meaning and may therefore be faster. We know that part of the process of L2 learning is to automatize some of those processes, so we can expect that if L2 learners are not familiar with certain conventional expressions, processing speed will be slower. Yet, research (Ervin-Tripp et al., 1987; Kasper, 1984) also shows that familiar contexts may help with nonconventional expressions. The impact of culture on one's ability to interpret implicatures is also important, because we know that part of the conventionality of implicature is anchored in the specific culture in which it belongs. Several examples of these differences in cultural conventions have been studied by various authors: jocularity of American sarcasm (Wallmark, 1971), use of honorifics in Japanese or Chinese (Brown & Levinson, 1987; Thomas, 1995), or even the fact that Malagasians are known to be vague

(Keenan, 1976). In the context of L2 processing, being unfamiliar with these cultural conventions can therefore hinder comprehension, unless L1 and L2 share the same conventions. In the event that the conventions (whether cultural conventions of communication patterns or conventional expressions) are unknown or too opaque, L2 learners may be able to rely on other contextual cues, as we have seen in our section of irony ([section 2.6](#)), which involve many different strategies (processing of visual or auditory cues such as gesture, intonation, using background knowledge or logical reasoning), which we will discuss in the next section ([3.3](#)). Despite evidence that learners can and do make use of multiple cues (even in situations in which L2 linguistic conventions are unfamiliar to them), Taguchi (2005, 2007, 2008b, 2008c, 2009a, 2009b, 2011) consistently argues that, since nonconventional indirect meaning is more opaque, it will take longer to comprehend due to increasing processing load (cf. Relevance Theory, Sperber & Wilson, 1995; Hamblin & Gibbs, 2003).

We have referenced Taguchi's experimental studies many times in this section and the previous one, but have yet to discuss her findings with regard to L2 comprehension of indirect meaning as it relates to conventionality. Let us now turn our attention to some of her papers (2005, 2007, 2008a, 2008b, 2008c, 2009a, 2009b, 2011—we are purposely excluding Taguchi et al., 2016, from this section since we will review that paper in [Chapter 4](#)) and see how she classifies indirect meaning in terms of conventionality and whether L2 learners had more difficulties processing some categories of indirect meaning. Drawing on Sperber and Wilson's (1991, 1995) Relevance Theory, Taguchi posits that more conventional implicatures should be easier and faster to comprehend because conventional features reduce the need for additional processing:

[S]ome implicatures are strongly conveyed, while others are weakly understood, due to the number of contextual cues that must be processed for comprehension. The greater the number of cues to be processed, the more extensive the search for meaning becomes, resulting in greater processing effort. When implicatures convey conventional meaning, that is, when speaker intentions are linguistically coded or embedded within predictable, fixed patterns of discourse, the listener may not attend to such contextual cues as background knowledge, mutual understanding, or flow of discourse, consequently reducing the processing effort.

Taguchi, 2005, p. 545

While this is certainly true in most situations, we argue that it is not necessarily the presence of multiple cues that makes the processing effort greater but the potential lack of saliency of

contextual cues, or the opacity of the implicature. For L2 learners, this may be particularly important as they may rely on a specific salient contextual cue (e.g., facial gesture) to reach the proper interpretation, without necessarily processing the rest of the cues, or even the literal meaning of the utterance (which could be too opaque if their level of proficiency is too low to comprehend the utterance on a linguistic level).

Across her various studies, she specifically classifies her indirect (or implied) meaning types as follow:

- **Conventional Implicatures** or **More Conventional Implicatures (MCI)** include:

- *indirect refusals*<sup>24</sup> (2005, 2007, 2008a, 2008b, 2008c, 2009a, 2009b, 2011)
- *indirect requests* (2005)
- *routines* (2009b, 2011)
- *conventional indirect opinions* (2008a, 2009a as well as Taguchi, Li and Liu, 2013)

- **Non-conventional Implicatures** or **Less Conventional Implicatures (LCI)** include :

- *indirect opinions* (negative and positive: 2005, 2007, 2008b, 2008c; “non-conventional indirect opinions”: 2008a, 2009b and Taguchi et al, 2013; or non-literal comments/indirect opinions: 2009b, 2011).

As stated previously, Taguchi’s conventional implicatures all follow a specific linguistic form or predictable language use pattern. Indirect refusals for example are considered conventional because it is a common language pattern “for one to provide a reason for refusing” (Taguchi, 2005, p. 517-548). Conventional indirect opinions only appeared in her studies that focus on L2 Japanese (2008a, 2009a) and L2 Chinese (Taguchi et al. 2013) because both languages have linguistic expressions that signal indirect opinions.

On the other hand, the non-conventional implicatures were found to be “more idiosyncratic and less stable across language users and contexts” (Taguchi, 2005, p. 548). Taguchi repeatedly considers indirect opinions as non-conventional because “they did not attach meaning to specific linguistic expressions or predictable patterns (e.g., indicating a negative opinion of a movie by saying “I was glad when it was over”)” (Taguchi, 2009b, p. 739). This classification somewhat departs from the definition Taguchi (2002) provides of indirect opinions (p. 166) and

---

<sup>24</sup> In Taguchi (2007, 2008b, 2008c) she specifies that indirect refusals include refusals to requests, refusals to invitations, refusals to suggestions and refusals to offers

their reliance on the adjacency pair rule (which we will talk about in our next section).<sup>25</sup> It is true that in English, one may have a variety of ways to expression an evaluative judgement, yet, we may wonder if indirect opinions that come as a reply to a request for evaluative judgement do not in fact follow some sort of conversation pattern as we have seen in [section 2.7](#). With the exception of Taguchi (2009b), all other studies by Taguchi that we reviewed (2002, 2005, 2007, 2008, 2009a) illustrate indirect opinions with an example of this conversational pattern.<sup>26</sup> Even Bouton (1994b) stated that indirect criticisms are semantically formulaic,<sup>27</sup> and as such, we could question Taguchi's classification of indirect opinions as non-conventional.

Results across all the previously listed studies from Taguchi were fairly consistent: implicatures that were based on a conventional form or pattern (in particular indirect refusals) were easier to understand for L2 learners. L2 learners took overall longer and were less accurate with LCIs or non-conventional implicatures. Indirect opinions were more challenging: L2 learners were slower and less accurate in their interpretation. According to Taguchi, these results support the claim that conventionality impacts comprehension of conversational implicatures:

Familiarity with these conventions reduces the processing load on the listener as well, allowing the listener to spend processing resources on conversational information that is genuinely new. When features of conventionality are not used by the speaker, more processing effort is required by the listener. The listener then requires more time to draw inferences about the speaker's intended meaning and the implications for the pragmatic context (i.e., how the speaker is representing information about the conversation topic).

Taguchi, 2005, p. 545

In this section we have shown that categorizing implicatures or types of indirect meaning is not always an easy task. Whether authors (Bouton, 1988, 1992, 1994a, 1994b; Yamanaka, 2003)

---

<sup>25</sup> Taguchi (2002), p. 166 "Because the adjacency pair rule is normative, the violation of the rule became salient incidence for the learners during listening. A question asking for evaluating comments should be followed by "good" or "bad.""

<sup>26</sup> For example, Taguchi (2005), Table 3, p. 549 provides two examples of LCIs: the Negative Implied Meaning example follows a conversation in which John asks Mary "How was the wedding? I bet it was exciting?" and Mary replies "Well... the cake was OK." The nonnegative Implied Meaning example shows Dave asking Susan if she likes the people upstairs, to which she replies "We're always visiting each other."

<sup>27</sup> Bouton (1994b, p. 99) on defining indirect criticism: "there is no obvious structural formula but there is a semantic one [...]. In this case, the speaker often responds with a positive remark about some peripheral, unimportant feature of whatever (s)he is asked to evaluate."

use Grice's Maxims or categorize indirect meaning in terms of speech acts or tropes (Cook & Liddicoat, 2002; Kim, 2013, 2014, 2016; Taguchi, 2002, 2005, etc.; Takahashi & Roitblat, 1994), the definition and limits of each categories are not always clear and there is sometimes an overlap between them. It is therefore hard to identify specific types of implicatures or speech acts that may be challenging. Based on the research we reviewed, we believe it is however reasonable to argue that conventionality—or L2 learners' knowledge and familiarity of the conventions (whether linguistic, cultural, or pragmatic)—can facilitate indirect meaning comprehension.

In [Chapter 4](#), we will discuss Taguchi et al.'s (2016) paper and how its findings are not consistent with Taguchi's previous findings with regard to conventionality. But before we get to this, we wanted to quickly review existing literature that addresses how L2 learners are reaching their interpretation. What strategies do they use to comprehend implicatures and how might these strategies be influenced by their proficiency, the type of implicatures they are processing, and the context in which they are produced?

### 3.3. Inferencing Strategies used for L2 comprehension

Strategies used to reach the interpretation of implicatures may be based on several factors. People may use different strategies depending on the types of implicatures they are processing and whether they are conventional or not. Additionally, proficiency may also impact strategy selection. Lower-proficiency learners may rely on certain strategies to compensate for their lack of cultural or linguistic knowledge. That may be true, even in the case of conventional implicatures, if the learners are not familiar with the conventions on which the implicatures are based. These strategies often rely on various types of contextual sources that we have discussed when we reviewed theories focusing on irony (Attardo et al., 2013; Attardo & Poggi, 2003; D'Imperio et al., 2013; Gibbs, 2000; Haiman, 1998; Shively, 2013; Wallmark, 1971; Yus, 2000, 2017). We already know from Yus's (2000) theory on optimal accessibility, that multiple contextual sources can be activated at the same time. However, Taguchi (2002, 2005, 2009a, 2009b) argues that when conventional implicatures are used, meaning is processed faster and therefore people do not have to rely on contextual cues as much. In line with Yus (1998, 2000), we believe that it is the lack of saliency of a specific convention, or a specific cue, that actually leads a hearer to look for additional cues. This phenomenon is one that has already



been explained by several authors (Attardo 2000a, 2000b, 2002, Giora 1995, 2003; Yus, 2000, 2017, etc.) and that is relevant for L2 comprehension since learners may not be familiar with certain conventions, which means that the conventions would not be “salient” to them. We may remember that according to Wallmark (1971), Sperber and Wilson (1995), Thomas (1995), or Yus (2017) (to only reference a few), contextual cues can derive from the following:

- *encyclopedic or factual information*: this includes general background knowledge, prototypical and situational expectations. In essence, these rely on certain conventions of conversation patterns or cultural norms.
- *specific situation or context*: unlike the above, this one may vary based on the specific physical environment in which the conversation takes place.
- *mutual knowledge*: assumptions the hearer and speakers may have about each other, their personality, preferences, etc.—these are not culturally conventional but may create an expectation of the way the specific people engaged in a conversation will act and react, which may impact comprehension. For example, if a person is talking to someone who is particularly sarcastic (and the person knows that), they will expect them to use sarcastic comment and utterance processing may be done in light of that fact.
- *phonological marker*: specific intonation, emphasis, pauses, laughter, etc.—these could be conventional or not.
- *kinesic markers*: body language such as gestures, facial expressions, etc.—some of these may be universal and understood by anyone but some may be culture-specific. Some gestures may also convey a specific attitude from the speaker (which will be conventionally accepted) and some may be more idiosyncratic and only apply to a specific situation.
- *linguistic markers*: specific syntactical or lexical devices—these can be conventional expressions that are known to convey a certain intention or illocutionary force but may also be situation-specific.
- *co-text*: the role of the previous utterances in the conversation. In the case of irony, it could be a reference to something previously said—cf. Sperber & Wilson’s echoic mention theory (1981) or Cutler’s provoked irony (1974)—or in the case of an indirect refusal, the fact that the person is replying to a preceding invitation. Though we see the reference to previous utterances as relevant in terms of meaning content, co-text can also be seen through the lens of Conversation Analysis (CA) (c.f. [section 2.7](#)). CA provides many examples illustrating the idea of preferred turn-taking (Sacks et al., 1974, Schegloff & Sacks, 1994; Schegloff, 2007; Kerbrat-Orecchioni, 2008; Taguchi, 2002) which is based on the linguistic context (Thomas, 1995), but often relies on cultural conventions. For

example, we know that the adjacency-pair rule (Schegloff, 2007; Schegloff & Sacks, 1974) relates consecutive utterances produced by two different people in an expected turn. Thomas (1995) illustrates this point with the example of Yes/No questions which “heavily constrain the possible interpretations of the replies” (p. 138). Indeed, when someone asks another person: “Will you marry me?” they expect their interlocutor to provide a positive/negative (or possibly neutral? maybe the person needs to think about it) answer. An indirect utterance (or even the lack of reply) would still be interpreted in light of this expectation. In the case of an indirect opinion, we may also argue that the co-text (the presence of a request for an opinion) can create a certain expectation: if someone is asking for someone else’s opinion, we should expect the interlocutor to provide some kind of evaluative judgment in their reply (positive, negative or neutral).

Very few studies have investigated the types of strategies L2 learners use to comprehend indirect meaning. Some have looked into the strategies used to produce indirect meaning, and how successful these strategies are, or have explored the teachability of implicature comprehension or production (and designed an instructional approach based on such strategies—cf. Bardovi-Harlig & Vellenga, 2012; Bouton, 1994b, 1999; Ishara & Cohen, 2010; Kim, 2016; Kim & Lantolf, 2016, who used awareness-raising activities<sup>28</sup> followed by pragmalinguistic discussions to ensure learners would notice, process and/or use the contextual cues we listed above). In this section, we will concentrate on studies that focused on the use of inferential strategies by learners during the comprehension process. These studies used verbal interviews to collect data. We will present our review of this literature in the form of a list of strategies that have emerged through the analyses conducted in these studies. We will also discuss how some strategies seemed to be favored by lower-proficiency or higher-proficiency learners or why some types of implicatures have been associated with the use of specific strategies.

- **Visual cues:** This category only appeared in studies that used a multi-modal instrument that includes videos (Yamanaka 2003; Shively et al., 2008; Kim, 2013; and Taguchi et al., 2016). Yamanaka (2003) does not actually present a systematic analysis of her participants’ strategies because the information they provided was inconsistent,<sup>29</sup> but she answers her third research

---

<sup>28</sup> Applying Schmidt’s (1998) Noticing Hypothesis to pragmatic phenomenon such as humor, sarcasm, etc.

<sup>29</sup> Yamanaka (2003) on explaining why she could not provide a quantitative analysis of the strategies used by her participants: “some provided clues for each vignette, some only for certain vignettes, and some none at all” (p. 118).

question (“Do Japanese learners who answer an item incorrectly on the questionnaire fail to notice contextual clues used by the NSs for that vignette?” p. 114) positively. She provides some examples of the contextual clues her participants used and whether they were successful in interpreting the implicature, among which (in her Appendix E, pp 166-175): gestures (“covers face”), body language (“looks away,” “fools around with props”), and facial expressions (“sneer”). She notes that while native speakers and many non-native speakers relied on these nonverbal language devices, some non-native speakers drew an incorrect interpretation from these visual cues and essentially misinterpreted the facial expressions or body language of the characters. Kim (2013) observed that her L2 participants relied on facial expressions at a much higher rate than the native speakers (who relied more on intention or context) to interpret sarcastic comments. She hypothesizes that her Korean participants may have had to rely on visual cues because they were not able to access other contextual sources (“factual information,” “biographical data,” “mutual knowledge,” etc.—cf. Kim, 2013, p. 195-196 as well as [sub-section 2.6.4](#) for a list of contextual cues). She also does not exclude the possibility that Koreans may use nonverbal cues to a lesser extent than American speakers, which could render facial expressions much more noticeable for the Korean participants. Kim also found that L2 participants were noticing certain “body movements” (p. 200) and gestures, and relied on them to identify sarcasm. On the other hand, Shively et al. (2008) found that lower-proficiency learners did not attend to visual cues when they were available because they were focusing on sentence-level meaning, and the bottom-up processing effort that was required for the comprehension of linguistic meaning prevented them from processing visual and auditory cues (when comparing groups who only had a written test versus groups that had an audio-visual instrument, no significant differences were observed).

- **Auditory cues:** these paralinguistic clues are defined by Taguchi (2002, 2009a) as being “used by the speaker to show his/her attitudinal and emotional state toward the message” (2002, p. 159). They include “tone of voice” (Taguchi 2009a, p. 266, and Yamanaka, 2003, p. 166, 167, etc.), “intonation, stress, pause, tone or speech rate” (Taguchi, 2002, p. 159, Taguchi et al., 2016, p. 690), and Taguchi (2002, 2009a) reports that these cues were used at the same frequency rate by her low- and high-proficiency learners. Yamanaka (2003) notes that reliance on auditory cues was not always successful however, as some non-native speakers misinterpreted the intention or attitude of the speaker based on their tone of voice. Just like with the visual cues, Taguchi (2002) also found that auditory cues were sometimes useful for learners to reach the intended meaning of an utterance, even though they may not have full access the linguistic meaning of the sentence (p. 165):

Paralinguistic features served as salient cues in understanding the attitudinal/emotional aspects of a message, and helped learners draw inferences about the speaker's intended meaning, particularly when they could not comprehend the language.

- **Key word inferencing:** this strategy involves “catching a key word in the input and associating it with meaning” (Taguchi, 2002, p. 159). It seems to be used primarily by low-proficiency learners. It relies heavily on linguistic encoding (primarily semantic) but is often the sign of poor sentence-level comprehension. Indeed, when learners struggle to understand the full dialogue, they are more likely to focus on specific word they understand (Ross, 1997; Taguchi, 2002). Garcia (2004) notes that in her study, learners' comprehension of certain speech acts was facilitated by the recognition and identification of specific linguistic or lexical cues. However, this strategy denotes a very basic bottom-up approach that often leads to an incorrect interpretation. Taguchi (2002, or Taguchi et al., 2016) reports that some learners “relied on the wrong key word, which led to a wrong reference” (2016, p. 690). Sometimes they may even hear a word that was not in the dialogue due to their poor listening skills (Taguchi, 2009a). Yamanaka (2003) observed the same phenomenon: some of her examples of unsuccessful contextual clue uses include references to specific words in the dialogue that were misinterpreted by the non-native speakers. While not strictly a keyword inferencing strategy, Cook and Liddicoat postulate that low-proficiency learners who are unfamiliar with the target language conventions, rely on “word-by-word analysis” (p. 32) at the expense of top-down processes that would rely on non-linguistic context cues. Such an attention to the linguistic meaning of the utterance may also prevent low-proficiency learners from using pragmalinguistic strategies and attending to other contextual cues.

- **Background knowledge/Experience:** Taguchi (2002) defines this strategy as “referring to life experience and global background knowledge” (p. 159). The examples provided in her study seem to refer to the use of the learners' personal experience to interpret the implicatures or relate to the situation on which it is based: “If I get good grades, I would be happier and I would show my grades right away” (2002, p. 159). There were some instances in Taguchi's classification where strategies seem to overlap and the distinction between background knowledge/experience and other categories like logical reasoning or speaker intention was not always clear. Low-proficiency learners in Taguchi (2002) relied on background knowledge more often than high-proficiency learners (a result that was inconsistent with Taguchi et al., 2016). While Yamanaka (2003) did not label any contextual clues as background knowledge or experience (she often just used “others” as a classification for clues that did not strictly rely on paralinguistic cues, or reference to the implicature itself), there are a few examples in her paper

that seem to illustrate that a non-native speaker was indeed using their personal experience to interpret the implicature. She actually reports that on one instance, a non-native speaker “may have transferred her own feelings toward [the characters]” (p. 136). Because life experience and of course “global knowledge” do not happen outside of culture, it seems to us that this category of strategy can also be associated to cultural conventions in the sense that experiences are necessarily culturally (and individually) specific.

- ***Adjacency Pair Rule***: More specific than the previous category as it strictly relates to pattern of conversation, this strategy clearly relies on the conventions of communicative exchange. Taguchi (2002) actually references Conversation Analysis in her definition and indicates that learners will rely on the specific sequence of utterances and how “the parts are dependent upon one another for coherence” (p. 159). We believe that specific types of implicatures (like indirect refusals, indirect replies, and even indirect opinions) are usually primed by the previous utterance and therefore this contextual cue is always present whether learners use them or not (and whether they actually verbalize their process in such a way that includes this rule—they may have actually been aware of the expectation the previous utterance creates but not explicitly mention it in their verbal reports). In some instances, the implicature may be noticeable because it violates the adjacency pair rule. Taguchi (2002) reports that L2 learners recognized the nonconventional or unpreferred turn in disclosures, due to the fact that the utterance did not answer the preceding request for information (in the example she provides, the speaker did not want to disclose incriminating/embarrassing information and admit to being arrested for being drunk). A second example of this phenomenon appears in Taguchi (2002) when a learner recognizes that the character did not respond directly to a request for an opinion (in the scenario, a student asks her professor what he thought about her presentation and he provides an indirect opinion “It’s hard to give a good presentation” p. 163). We should note that the fact that Taguchi considers indirect opinions to be a violation of the adjacency rule seems to imply that indirect opinions can follow a conventional pattern (since she says “the adjacency pair rule is normative,” 2002, p. 166). Yet, in all her studies, she views indirect opinions as nonconventional and idiosyncratic in nature.

As far as whether proficiency affects the use of this strategy, Taguchi (2002) reports that it was one of the most used (along with paralinguistic cues) by both the low and high-proficiency learners, but that the former tended to use it more than the latter.

- ***Speaker Intention***: This strategy only appears in Taguchi (2002), at least in this specific categorization (it looks like some of the participants in Yamanaka (2003) may have also verbalized strategies based on the recognition of the speaker’s intention). Taguchi defines it as

“understanding the function of the implicature spoken by the speaker; why and for what purpose the speaker used the implicature instead of a literal response” (p. 159). This category relies on the explicit verbalization by the participant of the utterance illocutionary force. In Taguchi (2002) it often appears alongside other strategies (paralinguistic cues, adjacency pair rule). It is not clear however how this strategy might differ from logical reasoning since all the examples provided by Taguchi seem to involve some type of reasoning explaining the mechanisms behind the implicature (and maybe it was included in the logical reasoning category in later studies). It is also possible that this strategy was particularly relevant in the processing of disclosure (a type of implicature Taguchi used in her 2002 paper but not in later studies). This strategy along with paralinguistic cues and the adjacency pair rule was used more frequently by higher proficiency learners.

- **Logical reasoning:** Taguchi (2002) provides a definition for this strategy as “recognizing literal meaning and working deductively toward the implied meaning.” In her study, both low and high proficiency groups were found to use logical reasoning at the same frequency rate. If Taguchi’s (2002) logical reasoning is based on the learners’ ability to verbalize the implicature mechanisms, we wonder whether references to specific conventional patterns would be included in this strategy, or if this type of strategy focuses more on nonconventional implicatures that do not rely on specific linguistic features or implicature that may not present other types of contextual markers (such a visual or auditory cues). More problematic for us is the fact that, when it comes to indirect meaning comprehension, we speculate that any inferencing strategy will rely on logical reasoning to some extent (whether the hearer verbalizes that process or not).

- **Situational/contextual cues:** There were some cues that participants in both Yamanaka’s (2003) and Taguchi’s (2002) studies detected in the video clips that did not strictly belong in the visual cues as described above (based on Taguchi’s definition). Yamanaka referred to them as “other visual cues” or “other audio clues.” We refer to them as situational cues because they could be anything that may relate to the setting of the video: a character’s “appearance,” the décor, the “music,” or a “scary alley” (Yamanaka, 2003, p. 167), props, etc., that learners may take into account when interpreting a context-based implicature. This strategy does not include visual and auditory cues previously listed as these relate strictly to the speaker. The hearer’s reaction (body language) could also be included in this strategy (Yamanaka, 2003). A few of Yamanaka’s participants used these situation-based clues with a mixed level of success.

- **Test-taking strategy:** elimination process. Taguchi (2002), Yamanaka (2003) and Taguchi et al.’s (2016) studies rely on a multiple-choice questionnaire, but only Yamanaka references her

participants basing their answers to the questionnaire on a strategy involving an elimination process. When presented with four options, participants explained that they did not select specific answers and may have narrowed down their options or selected one answer by default after eliminating all the others. This strategy of course is different from the other ones as it is specific to the study and instrument design. In Yamanaka's (2003) data, it was often used in combination with other strategies. For example, a learner may eliminate one option because they did not hear a specific word referenced in that option, or did not note any negative facial expressions that may have been conveyed by that option, thus combining several of the above strategies with an elimination process strategy.

We have seen above that some of the strategies presented by the authors we reviewed are very clear and easy to identify while others may have more problematic definitions. Some studies also show a clear difference in the use of strategies by high and low proficiency learners or a more frequent reliance of a specific strategy for one type of implicatures. We have mentioned that some strategies were often combined with other strategies. This multiple activation of contextual cues has been addressed before, especially with the processing of irony comprehension. It is worth noting that the authors we discussed in this section often discuss the relationship between the use or need to rely on multiple cues and proficiency or degree of opacity/conventionality. Taguchi (2002) notes that learners did not use multiple strategies to comprehend indirect refusals because "they were embedded in conventional, easily recognizable contexts" (p. 162). She also reported that both lower-proficiency and higher-proficiency learners used a similar number of strategies overall but lower proficiency students used background knowledge/life experience as well as key word inferencing more often, whereas more proficient students were able to identify the speaker's intentions. This led Taguchi to conclude that "learners' different experiences and ability with language might shape their context differently and provide different access to certain types of information" (p. 170). Unlike Cook and Liddicoat (2002), findings from Taguchi's (2002, 2005), Garcia's (2004) and Yamanaka's (2003) support the claim that "learners were able to interpret implicit message even when they did not completely understand the language in the test sentences. Learners used contextual cues which were salient for interpretation" (Taguchi, 2002, p. 164). Nevertheless, lack of knowledge of cultural norms or linguistic conventions does increase the opacity of implicatures for learners who will then have to turn to additional contextual cues: "Although the learners relied on other strategies (e.g., tone of voice) to compensate for the lack of

pragmalinguistic knowledge, unfamiliarity of the forms seemed to have posed a challenge during comprehension” (Taguchi, 2009a, p. 266).

Our next chapter ([Chapter 4](#)) will be devoted to a review of Taguchi et al. (2016). We excluded this paper from the previous literature review because we wanted to carry out a more in-depth analysis of its design, methodology and findings, since this dissertation aims to replicate this study. We hope to not only comment on Taguchi et al.’s results, but also reflect on some of their findings and limitations, as well as bring to light some potential issues with the methods or conclusions.



## Chapter 4 - A close look at Taguchi et al.'s (2016) study

In 2016, Taguchi, Gomez-Laich and Arrufat-Marques published an article in the *Foreign Language Annals* entitled “Comprehension of Indirect Meaning in Spanish as a Foreign Language.” They report on a study that “investigated comprehension of indirect meaning among learners of L2 Spanish via an original computer-delivered multimedia listening test” (Taguchi et al., 2016, p. 677). As such, the study expands upon previous studies from Taguchi, which we have reviewed in [Chapter 3](#) (Cf. [Appendix A](#) for a summary of these studies which provides an overview of their differences and similarities). The purpose of the 2016 study was to “examin[e] theoretical claims about inferential mechanisms and their applicability to L2 comprehension.” (p. 677). They selected three different types of indirect meaning, based on their degree of conventionality, and the amount of inference needed to interpret them: indirect refusals, irony, and indirect opinions. Their overarching research question was “Do L2 Spanish learners demonstrate different comprehension accuracy and response speeds across different types of indirect meaning?” (p.685), which addressed the following four areas of study as highlighted by Taguchi et al. (p. 683/my numbering):

- 1) to what extent learners are able to comprehend indirect meaning,
- 2) what causes comprehension difficulty,
- 3) what strategies they use to draw inferences,
- 4) the effectiveness of an original multimedia instrument in assessing L2 learners’ comprehension accuracy and response speed for different item types representing different kinds of indirect meaning in Spanish.

They used a mixed method research methodology: a quantitative analysis of the comprehension scores and response time was performed, while a more qualitative approach was used to analyze the findings from retrospective verbal interviews investigating the difficulties encountered and strategies used by a subset of participants.

In this chapter, we will provide a detailed analysis of the study Taguchi et al. published, including how they defined the three indirect meaning categories above, and offer a critique of some elements that were problematic for our replication.

## 4.1. Categories of indirect meaning

As mentioned previously, Taguchi et al. decided to focus on three different types of indirect meaning for this study: indirect refusals, indirect opinions and irony. The first two types of implicatures had been studied by Taguchi before (2002, 2005, 2007, 2008, 2009a, 2009b, 2011), but this was the first time Taguchi included irony as a category. We have seen in [Chapters 2](#) and [3](#) that irony is not always easy to define. Other authors (Bouton, 1988, 1992, 1994a, 1994b, 1999; Bromberek-Dyzman & Ewert, 2010; Shively, 2013; Shively et al. 2008; Yamanaka, 2003) have looked into L2 comprehension on irony, but the difference with respect to other categories used by these authors (such as sarcasm, or even Pope Q) was not always clear (Yamanaka, 2003).

Taguchi et al.'s central argument (one that Taguchi made in previous studies as well) is that difficulty in comprehension stems from the degree of indirectness of the implicatures, and particularly the degree of conventionality of the implied meaning utterances. They state that “comprehension is easier and faster for conventional implicatures than nonconventional ones, as long as the conventionality encoded in the utterance is shared between L1 and L2” (p. 679). In what follows, we will begin by explaining how Taguchi et al. operationalized their three categories of indirect meaning and how they relate each to the concept of conventionality. We will then turn our attention to the challenges that such a categorization presented in their analysis and for our replication.

### 4.1.1. *Indirect refusals*

In Taguchi et al. (and other papers we reviewed), indirect refusals were utterances that came as a response to a request, an invitation or a suggestion, but that did not use direct and explicit expressions (IFIDs—see [section 2.7](#)) such as “No,” “I can’t,” “I don’t want to.” Taguchi et al. explain that expressing refusals indirectly (without the use of an IFID) is a common pattern that “reflect[s] the conventionality of language use” based on a “specific discourse pattern” (p. 684) of responding by providing a reason for the refusal. While this type of conventional pattern is common in many languages (Taguchi 2008, 2009a, 2011, Taguchi et al., 2013), differences can sometimes be found between cultures. For example, Chen, Ye, & Zhang (1995) showed that a conventional pattern of indirect refusal in Chinese combines two strategies: providing a reason and offering an alternative. In the case of Taguchi et al.'s study however, L1 English and L2

Spanish conventions are similar (Felix-Brasdefer, 2008). Taguchi et al. provide the following example of indirect refusal used for their study (we only use the English translation below but the Spanish version can be found on p. 685):

*Antonio:* Hi, Margarita! I ordered something for you to drink. I ordered you a hot chocolate. It's so cold today.

*Margarita:* Thank you, Antonio!

*Antonio:* Do you want something to eat? I didn't order anything because I didn't know what you'd like.

*Margarita:* I had lunch at home.

In the above example, Margarita does not explicitly decline Antonio's offer to order some food. The intended meaning of her utterance is that if she already had lunch, she is likely no longer hungry or does not want to eat any more, therefore indirectly refusing Antonio's invitation to order food. Yet, while we agree that providing a reason or explanation for a refusal is common practice (following theories of Politeness), we know that IFIDs can co-occur with this convention (providing a reason). We do not know, however, how often indirect refusals occur in the total absence of an IFID. This means that we do not know how natural this is in actual conversation, or how often language users are required to interpret indirect refusals without being "aided" by an accompanying direct refusal indicator.

#### 4.1.2. *Indirect opinions*

Indirect opinions were said to be less conventional than indirect refusals (and irony) in the sense that they did not present a specific pattern of response ("linguistic options for expressing the opinion [...] are wide open and more idiosyncratic and free than those of the indirect refusals used in this study," 2016, p. 684) and that a multitude of possible answers could be imagined (each following a different and unpredictable pattern). Taguchi et al. presented the following example, in which two friends are talking about a wedding they attended. The last two turns include a question from the male speaker (a request for opinion) and an indirect opinion from the female speaker:

*Rogelio:* And did you like her dress?

*Ana:* If we were in the 80s.

Excerpt from Sample Listening Test Items, Taguchi et al. (2016), p. 685.

The intended meaning in Ana's utterance is that she did not like the dress. However, such indirect opinion can also be said to express indirect negative criticism using sarcastic irony. The

above example would therefore work along the same lines as other ironic criticisms (or diasym) as defined by many authors in both English and French (Berrendonner, 1981; Bouton, 1988, 1990, 1994a, 1994b; Charaudeau, 2006, 2011; Clark & Gerrig, 1984; Colston, 2000; Gibbs, 1993, 1994, 2000, 2002; Kumon-Nakamura, et al. 1995; Sperber & Wilson, 1981; etc.). In their paper, Taguchi et al. do not specify whether the indirect opinions items focused on negative or positive evaluations and whether that aspect may impact comprehension.<sup>30</sup> Regarding conventionality, we have, in [Chapter 3](#), discussed the fact that some authors (e.g., Bouton or even Taguchi, 2002) suggest that some indirect opinions may actually follow a predictable pattern of conversation as they often come directly after a request for an evaluative judgement. Therefore, in the above example the male speaker's question ("did you like her dress?") may set an expectation for the next turn and the female speaker's contribution may therefore be interpreted in light of this expectation. Conversation Analysis explains that this adjacency pair rule primes the answer and even in the presence of an indirect meaning utterance, people will still expect it to be some sort of positive, negative or potentially neutral evaluation.

#### *4.1.3. Irony*

Taguchi et al.'s rationale for including irony in their study originates from a need to expand the scope of indirect meaning types (since Taguchi had not previously focused on this type of implicature before) and to add to the very limited research that had been done on the L2 comprehension of irony. Drawing on Grice's maxims, Taguchi et al. explain that "irony is a type of conversational implicature that flouts the maxim of quality" (p. 681). Using the traditional neo-Gricean definition of irony (an utterance that is used to mean "the opposite of what was actually said," p. 684), Taguchi et al. state that irony is a common conventional rhetoric device, despite presenting a "greater deviation from the literal meaning" (p. 681) and thus being even more indirect than indirect refusals and indirect opinions.

Irony items in Taguchi et al.'s study fell in two categories: "meaning reversal" or "meaning replacement." Quoting Kapogianni (2014), they explained that "meaning reversal" irony involves a process by which the speaker says the opposite of what they mean by using a semantically related utterance (example provided p. 681: "saying "I love grading papers" to

---

<sup>30</sup> However, in the document that Taguchi sent us, upon our request, with all the indirect meaning tasks they used, all indirect opinion items were offering some sort of negative evaluation. In previous papers (2005, 2007, 2008), Taguchi had included a mix of positive and negative indirect opinions<sup>30</sup> though she never really analyzed the differences in comprehension within this group.

mean “I hate grading papers,”” Taguchi et al., 2016). On the other hand, with “meaning replacement,” the ironic comment has no semantic relationship with the intended meaning; rather “the intended meaning replaces the proposition” (Taguchi et al., 2016, p. 681). They provided an example from Kapogianni (2014) in which “B’s utterance is meant to be the opposite of the proposition but it is not semantically related to A’s utterance” (p. 682):

A: Was I speeding?

B: No, you were standing still. It’s the scenery that’s been whizzing by at 90 mph.

(Kapogianni, 2014, p. 599)

This typical pattern of opposition—whether meaning reversal or replacement—is the reason why Taguchi et al. classified their category of irony as conventional. It is easy to see why Taguchi et al. may have wanted to restrict their irony types to two that followed a specific pattern (meaning reversal or replacement): given that they were interested in investigating the impact of conventionality on L2 (irony) comprehension, they needed to select irony examples that were based on a “common rhetorical device” (p. 681). Yet, we will admit that what Taguchi et al. refer to as “meaning replacement” (based on Kapogianni, 2014) does not seem to follow a very predictable pattern of formation to us. As discussed in [Chapter 2](#), the neo-Gricean definition of irony includes a wide array of idiosyncratic rhetorical devices. While we agree that meaning reversal is conventional, categorizing meaning replacement as conventional on the other hand seems like a stretch.

While Taguchi et al. stated their desire to “expand [their] understanding of L2 irony comprehension” by “investigat[ing] the two types of irony items (meaning reversal and replacement),” there is no report of analysis between the two categories in their paper. It is not clear whether the irony test items were coded to differentiate between reversal and replacement since Taguchi et al. did not include any analysis on the difference between the two types of irony in their study.

On a final note regarding Taguchi et al.’s irony items: all but one item (IRO6 which is an example of ironic praise) were sarcastic comment or indirect criticism (which is fairly representative of the natural distribution as we know negative irony is more common than positive irony).

#### 4.1.4. Potential challenges with the three categories

We already mentioned some of the issues with the three different types of indirect meaning Taguchi et al. used, in particular the problematic definitions of indirect opinions and irony and how they relate to conventionality. In order to replicate this study and to gain a better understanding of the type of tasks learners had to complete, we reached out to Taguchi to ask if she would agree to send us a copy of their instrument. Taguchi sent us a document that included all the tasks used in Taguchi et al.'s pilot study. A careful review of all the dialogues showed that at times it was hard to distinguish among the three types of indirect meaning categories (even though the dialogues were classified by types of indirect meaning). While all the indirect refusal items seemed to present a consistent and clear adherence to the definition for that category, items in the indirect opinions and irony were sometimes hard to separate or clearly label. For example, one could argue that the following dialogue, used by Taguchi et al. as an example of irony, could also be considered an indirect refusal as well as an indirect opinion:

This is a conversation between two roommates, Abel (male speaker) and Elsa (female speaker). They are talking about watching a TV show together.

*Abel:* Elsa, do you have plans for tonight?

*Elsa:* No, why?

*Abel:* Because there's a new TV show that I'd like to watch. Do you want to join me?

*Elsa:* You know me. I find TV really educational. Every time someone switches it on, I go to my room and read.

Elsa's last comment was problematic on several levels. First, the ironic comment "I find TV really educational" is followed by a literal comment that reflects the reality of Elsa's behavior which is not ironic in itself and actually has a direct meaning. We feel this could create some comprehension issues for L2 learners, especially since the indirect meaning (target sentence) was in the last sentence of the dialogues for all the other tasks. Even more puzzling is the fact that Elsa's comment could technically fall under all three categories of indirect meaning: while clearly ironic, she also answers Abel's question ("Do you want to join me") negatively in an indirect manner (i.e., indirectly declines his invitation), and also expresses an indirect opinion regarding TV shows (or TV in general and the activity of watching TV). The difficulty seems to lie with the fact that while both indirect refusals and indirect opinions are speech acts (which can be expressed directly as well), irony is not a speech act and can actually be used in

conjunction with many different speech acts. In other words, irony items in Taguchi et al. always fulfill some kind of speech act (expressing disapproval, giving an opinion, refusing an offer, etc.), but in several dialogues, irony comments seemed to offer a sarcastic indirect opinion (solicited or not). Similarly, some indirect opinions used by Taguchi et al. would be considered ironic or sarcastic.<sup>31</sup> Given the difficulty we had in separating the categories of indirect meaning for some of the items—categories that are supposed to trigger different patterns of inferences according to Taguchi et al.—we may wonder how the overlap may affect the analysis of L2 learners' comprehension as well as their use of inferential strategies.

On the subject of conventionality, Taguchi et al. had also said that indirect opinions did not follow any kind of conventional pattern. We have previously argued that some indirect opinions can follow a conventional pattern of conversation (derived from the adjacency pair rule). Among Taguchi et al.'s indirect opinion items, we believe that nine follow this specific pattern.<sup>32</sup> We also wonder how that may have impacted L2 learners' comprehension.

Now that we have discussed the categories of indirect meaning Taguchi et al. studied in their paper, let us turn our attention to their instrument.

## 4.2. Instrument development

For their study, Taguchi et al. developed a multimodal test which presented several advantages. First, they rightfully pointed out that previous studies mainly looked at learners' comprehension of non-literal language using written dialogues (referencing Bouton, 1990, 1992, 1994a, 1994b; Cook & Liddicoat, 2002, Roever, 2005) as representation of oral conversations. The issue with such a format is that it does not reflect natural oral language, and it does not give participants the opportunity to exploit paralinguistic cues to access meaning. We know that these paralinguistic cues (cf. Attardo 2000b, Attardo & Poggi, 2003, Attardo et al., 2013) can be critical to one's ability to comprehend non-literal language, therefore providing participants with not only auditory but also visual input was a very important and novel aspect of Taguchi et al.'s study. Preceding the publication of their article, Taguchi had already used auditory input in studies on L2 English (Taguchi, 2011), but she had noted that further development of audio

---

<sup>31</sup> We estimate that at least six indirect opinions are sarcastic.

<sup>32</sup> OPI 2, 4, 5, 6, 7, 8, 9, 10, and 13

tests was necessary, especially in other languages. Taguchi et al. (2016) also indicated their desire to investigate “the effectiveness of an original multimedia instrument in assessing L2 learners’ comprehension” (p. 683). To that effect, they wrote dialogues based on real-life conversations or examples they had come across in the media (movies or TV series, etc.). Additionally, they looked for occurrences of indirect meaning (specifically irony, indirect refusals and opinions) in corpora of spoken Spanish conversations (“*Corpus Oral de Referencia de la Lengua Española Contemporánea*, or CORLEC”). Since the multimodal test needed to be accessible to lower-level learners, they decided to modify and adapt the dialogues they had collected. In order to do so, they made sure that “most vocabulary and grammatical structures were taken from the elementary-level Spanish textbooks used in the institution” (p. 684). They provided a sample (see [Appendix B](#)) of the changes they made in the three categories in their paper. A comparison of the original dialogues and their adapted versions, as published in their paper, revealed only minor modifications that, in our opinion, would neither hinder comprehension nor take away from the impact of the indirect meaning or authenticity of the conversations. Taguchi et al. had two of the authors review the modified dialogues and ensure their appropriateness.

Each of Taguchi et al.’s dialogues were preceded by a brief description of the situation (in English for the L2 learners and in Spanish for the native speakers) and followed by multiple-choice questions (also in the participants’ native language) offering participants four randomized options: one correct answer and three distractors. The following principles were used to create the distractors (2016, p. 687):

- Principle 1:** The option contains a meaning opposite to the target indirect meaning.
- Principle 2:** The option contains words taken from the last part of the conversation
- Principle 3:** The option is related to the overall meaning of the conversation.

Additionally, Taguchi et al. noted a different degree of accuracy in the distractors as they relate to the correct answer (p. 687):

The first principle was established as the least plausible information. If participants understood the meaning of the indirect utterance, this option was the most remote choice. The second principle was established based on the recency effect— that people tend to recall the words they heard last (Atkinson & Shiffrin, 1971). The last principle was established based on the key word processing strategy. Previous findings have shown that less-proficient listeners tend to go directly to



the first association of the key word and the referent, while proficient listeners explore several potential associations before making a decision (Ross, 1997).

We would like to address some of the issues we noted with the multiple-choice questionnaire and the application of the three principles above. The description of the three principles that Taguchi et al. proposed (see above) seems problematic to us because Principle 2 might be based on the recency effect but clearly focuses on certain words from the dialogue (and for many items, the option based on that principle contained keywords from the rest of the dialogue as well). However, Taguchi et al. say that Principle 3 relies on keyword association, which we find confusing since this Principle is supposed to be based on the overall meaning of the dialogue. We wonder if keyword processing strategy (which Taguchi et al. later found was one of the strategies used by L2 learners<sup>33</sup>) may also be at work in Principle 2.

Another challenge we wanted to bring up is the fact that the L2 learners had to listen to dialogues in the target language (Spanish) but answer questions in English, which means that Principle 2 was harder to follow since it was based on the translation of the words in the dialogue (not the actual words in the dialogue). Let us look at the following example from the tasks Taguchi sent us (first example of indirect refusals used in Taguchi et al.).

---

<sup>33</sup> Taguchi et al. (2016) p. 690: “Key word inferencing: Picking up a key word and associating it with meaning.”

## Figure 1

Taguchi et al. (2016) – IRF1

IRF1cena	IRF1cena
<p>Esta es una conversación entre dos amigos, Carlos (hombre) y Adriana (mujer). Están haciendo planes para esta noche.</p> <p><i>Carlos: ¡Hola, Adriana! ¿Tienes planes para esta noche? ¿Te gustaría ir al cine hoy?</i></p> <p><i>Adriana: Pues no creo. Tengo mucho por hacer.</i></p> <p><i>Carlos: ¿Y qué tal ir a cenar? No nos llevará mucho tiempo.</i></p> <p><i>Adriana: Tú siempre tardas mucho en comer.</i> (38)</p>	<p>This is a conversation between Carlos (male speaker) and Adriana (female speaker). They are making plans for tonight.</p> <p>Carlos: Hi, Adriana! Do you have plans for tonight? Wanna go to the movies?</p> <p>Adriana: I don't think so. I have lots of things to do.</p> <p>Carlos: And what about getting dinner? It won't take us that long.</p> <p>Adriana: It always takes you forever to eat. (42)</p>
<ol style="list-style-type: none"><li>1. Adriana no quiere salir a cenar esta noche. (CORRECT)</li><li>2. Adriana va a ir a cenar con Carlos esta noche. (Principle 1)</li><li>3. Adriana piensa que en el restaurante <i>tardan</i> en traer la comida. (Principle 2)</li><li>4. Adriana piensa que tiene tiempo para cenar con Carlos. (Principle 3) (42)</li></ol>	<ol style="list-style-type: none"><li>1. Adriana doesn't want to go out for dinner tonight with Carlos. (CORRECT)</li><li>2. Adriana is going to have dinner with Carlos tonight. (Principle 1)</li><li>3. Adriana thinks waiters at the restaurant are <u>slow</u>. (Principle 2)</li><li>4. Adriana thinks she has time to have dinner with Carlos. (Principle 3) (38)</li></ol>

The Spanish version (Spanish dialogue and Spanish questionnaire) did not raise any issue and follows the principles: option #3 was reusing some words from the last sentence. However, in the English translation, the focus is on the word “slow,” which is not technically part of the dialogue. Two issues were therefore at play here: the use of L1 (English) in the MCQ questionnaire (which means that none of the words in the questionnaire were actually heard in the Spanish dialogues) and inappropriate translation equivalent (as it is the case in IRF1 where the Spanish verb “*tardas*” in the dialogue becomes the adjective “slow” in the English MCQ).

As mentioned above, the issue with keywords was found at another level: we noticed that sometimes keywords from the last sentence of the dialogue also appeared in other options of the multiple-choice questionnaire. Indirect opinion item #6 in Taguchi et al.'s study was presented as follow (Spanish and English version):

## Figure 2

Taguchi et al. (2016) – OPI6

<p><b>OPI6pelo</b>          Esta es una conversación entre dos amigos, Gerónimo (hombre) y Melisa (mujer). Están hablando sobre el nuevo corte de pelo de Melisa.</p> <p><i>Melisa: ¡Hola, Gerónimo! ¡Hace mucho que no te veía! ¡Luces muy bien!</i>  <i>Gerónimo: ¡Pues muchas gracias, Melisa! Veo que tú has cambiado el corte de pelo.</i>  <i>Melisa: Melisa: ¡Sí! ¡Me le corté la semana pasada! ¿Qué te parece mi nuevo estilo?</i>  <i>Gerónimo: Prefería el corte anterior.</i>          (41)</p> <ol style="list-style-type: none"> <li>1. A Gerónimo no le gusta el <i>corte</i> de pelo nuevo de Melisa. (CORRECT)</li> <li>2. Gerónimo cree que el <i>corte</i> nuevo le queda bien a Melisa. (Principle 1)</li> <li>3. Gerónimo acaba de cambiar su <i>corte</i> de pelo. (Principle 2)</li> <li>4. Gerónimo y Melisa fueron a la peluquería juntos la semana pasada. (Principle 3)          (42)</li> </ol>	<p><b>OPI6pelo</b>          This is a conversation between two friends, Gerónimo (male speaker) and Melisa (female speaker). They are talking about the new haircut Melisa got.</p> <p>Melisa: Hi, Gerónimo! Long time no see! You look great!          Gerónimo: Thank you very much, Melisa! I see you got a new haircut.          Melisa: Yes! I got my hair cut last week! What do you think about my new hair style?          Gerónimo: I preferred the previous haircut.          (43)</p> <ol style="list-style-type: none"> <li>1. Gerónimo does not like Melisa’s new <u>haircut</u>. (CORRECT)</li> <li>2. Gerónimo thinks that the <u>new haircut</u> looks good on Melisa. (Principle 1)</li> <li>3. Gerónimo just got a <u>new haircut</u>. (Principle 2)</li> <li>4. Gerónimo and Melisa went together to the hairdresser’s last week to get their <u>hair cut</u>. (Principle 3)          (38)</li> </ol>
---	--

In the above, of all the words that appear in the last sentence of the dialogue, “*corte*”/”haircut” is the only one that is used in option #3 (which follows Principle 2). However, we also see that the same keyword is used for option #1 and #2 in the Spanish version as well as all four options in the English versions. It would therefore be very difficult to analyze the selection of distractors because these options may not actually measure the recency effect as Taguchi et al. claimed.

Finally, there were some instances when one of the options that was meant to be a distractor, was not actually wrong per se. In the example below, option #3 (which follows Principle 2) is a paraphrase of the literal comment made by Estela.

### Figure 3

Taguchi et al. (2016) – IRF9

<p><b>IRF9torta</b>          Esta es una conversación entre dos colegas, Iván (hombre) y Estela (mujer). Están hablando de hacer planes para ir a probar una nueva torta en La Mansión del Chocolate.</p> <p><i>Iván: Estela, ¿Cuándo vamos a ir a probar la nueva torta en La Mansión del Chocolate?</i>  <i>Estela: Cuando quieras, Iván.</i>  <i>Iván: ¿Qué te parece si vamos mañana a las 5 de la tarde?</i>  <i>Estela: Tengo tickets para ir al cine a la sesión de las 5:15.</i>          (42)</p> <ol style="list-style-type: none"> <li>1. Estela no irá a probar la nueva torta con Iván mañana a las 5pm. (CORRECT)</li> <li>2. Estela e Iván visitarán La Mansión del Chocolate mañana a las 5pm. (Principle 1)</li> <li>3. Estela irá al cine esta tarde. (Principle 2)</li> <li>4. Estela e Iván probarán el café de caramelo juntos. (Principle 3)</li> </ol> <p>(41)</p>	<p><b>IRF9torta</b>          This is a conversation between two friends, Iván (male speaker) and Estela (female speaker). They are talking about making plans to go to The Chocolate Mansion to try their new cake.</p> <p><i>Iván: When are we going to The Chocolate Mansion to try their new cake?</i>  <i>Estela: Whenever you want, Iván.</i>  <i>Iván: What about going tomorrow at 5 in the afternoon?</i>  <i>Estela: I have tickets to go to the movies at 5:15pm.</i>          (37)</p> <ol style="list-style-type: none"> <li>1. Estela won't go to try the new cake with Iván tomorrow at 5PM. (CORRECT)</li> <li>2. Estela and Iván will visit The Chocolate Mansion tomorrow at 5PM. (Principle 1)</li> <li>3. Estela is going to the movies this afternoon. (Principle 2)</li> <li>4. Estela and Ivan will try the new caramel coffee together. (Principle 3)</li> </ol> <p>(42)</p>
---	---

In the above, option #3 is not representative of the illocutionary force of the utterance, which is the indirect refusal (i.e., Estela will not go the Chocolate Mansion with Ivan to try their new cake). However, since participants were asked to “select the correct answer based on the content of the conversation,” option #3 was not technically wrong.

While we wanted to point out the potential issues with Taguchi et al.’s multiple-choice questionnaire and associated principles, we should keep in mind that they did not actually conduct a post-hoc analysis of the selection of distractors in this particular paper (Taguchi had done so in her 2005 and 2008a publications). Therefore, the design problems highlighted above may not have a noteworthy impact on the published results. We will now continue our review of the development of the rest of their instrument.

Before creating the videos based on the dialogues they had written, Taguchi et al. gave four native speakers of Spanish a paper-and-pencil version of their test which included 16 experimental items per category (16x3=48) and 10 fillers that were randomized. “Accuracy

rates of their responses, along with plausibility of the conversations were confirmed” (p. 687) and additional modifications were made to some items based on the provided feedback. Video clips were then created using two Spanish native speakers. Based on their paper, it is not clear how and to what extent the dialogues were staged, and no information was provided regarding the type of instructions that were given to the actors. Length of videos was not reported in Taguchi et al.’s paper.

Taguchi et al. piloted the computerized audio-visual version of their Spanish test with 21 native Spanish speakers. After analyzing the participants’ answers and response time, they further modified the test, which was then reduced down to 36 experimental items (12 items per indirect meaning type, plus 10 fillers). They did that by “removing the items that had lower accuracy rates and/or longer response times than average” (2016, p. 687). In the end, Taguchi et al.’s final version of the test was as follow:

- “The test had 48 items: 2 practice, 10 fillers items, and 36 experimental items.” (p. 684)
- “Each conversation has between 38 and 42 words and included from four to five speaker turns.” (p. 684).
- “The [multiple-choice] options were given in English for the L2 Spanish learners because of different levels of Spanish reading ability in the L2 group.” (p. 686)
- “The total number of words in the multiple-choice option sentences was kept approximatively the same across all items (34-41 words in English; mean = 38.30,  $SD = 2.60$ )” (pp. 686-687).

### 4.3. Experiment

Taguchi et al. conducted their research at a private American University which offered Spanish L2 courses. Their study design and methods are summarized below.

#### 4.3.1. Test Participants

Thirty-two participants were recruited from intermediate (third semester) to advanced (sixth semester) level Spanish classes (21 females and 11 males, mean age=19.9, p. 683). All participants were either native English speakers (95%) or had a high level of English

proficiency. They also reported that a few students had limited study abroad experience in a Spanish-speaking environment (less than three months).

In order to confirm participants' level of L2 proficiency, the NCSSFL-ACTFL Can-Do test was used. This test consists of a series of "Can-Do" statements that participants must check based on a personal assessment of their own abilities. Taguchi et al. selected fifteen of these statements from the Intermediate High and Advanced Low levels (on the ACTFL scale),<sup>34</sup> based on their relation to the activities that participants completed during the experiment. Their average score on a five-point scale was 4.24 for the 15 items ( $SD = 0.37$ ; range = 3.40-4.93) (p. 683). According to Taguchi et al, the mean was consistent with the courses from which the participants were recruited and the levels they targeted for their study.

The pilot study was administered to a group of 21 native Spanish speakers from various Spanish-speaking countries who were studying in the U.S. They were recruited via email and their age ranged from 19 to 48 years old ( $M = 28.95$ ,  $SD = 5.62$ ).

#### 4.3.2. Survey administration

It is unclear whether the pilot study that was administered to the native speaker group in Taguchi et al.'s study occurred under the same conditions as the L2 speaker group. Indeed, Taguchi et al. indicated that their native speaker group was composed of students "studying in a private or public U.S. university." (p. 684), which led us to believe that they may not have taken the test under the same conditions as the learner group (all from the same private university).

For the experimental study, Taguchi et al. reported the following procedures used with their Spanish learners, who completed the survey in a computer lab on the researcher's campus (p. 687):

After sitting in front of the computer, participants put on headphones and read the instructions on the screen. Videos were played on the computer. Immediately following each conversation, a multiple-choice question appeared on the screen [...]. Participants responded by

---

<sup>34</sup> As a point of reference, the ACTFL Intermediate High and Advanced Low levels are usually considered equivalent to B1.1 and B1.2 levels of the Common European Framework of Reference for receptive skills

([https://www.actfl.org/sites/default/files/reports/Assigning\\_CEFR\\_Ratings\\_To\\_ACTFL\\_Assessments.pdf](https://www.actfl.org/sites/default/files/reports/Assigning_CEFR_Ratings_To_ACTFL_Assessments.pdf)). It is surprising that language students in a third semester of language instruction at an American university would reach B1 level. We usually expect these students' receptive skills to be at a level A2.1 upon completion of a third semester French course (which would rate at the Intermediate Low/Mid level of the CEFR scale).

pressing the corresponding number on the keyboard. Once they chose the answer, the computer automatically took them to the next item. Response time was measured between the moment when the question appeared on the screen and the moment when the participants pressed the number key.

This part of their study was designed to measure comprehension accuracy and speed.

#### *4.3.3. Retrospective interviews*

In order to investigate the types of strategies and processes L2 participants used to comprehend indirect meaning (“e.g., what strategies were used to infer meaning and which cues were attended to,” p. 683), Taguchi et al. conducted verbal interviews with a smaller group of participants. They divided L2 respondents into three groups: the top third corresponded to participants who received the highest scores on the listening test, and the bottom third included participants who scored the lowest on that test. Four participants were randomly selected from both the top and the bottom groups (with a total of eight students: seven females and one male). Taguchi et al. did not specify whether assessed proficiency level corresponded to low or high scores (i.e., whether the participants who scored the lowest were part of the intermediate proficiency group and participants who scored higher belonged to the more advanced courses with a higher proficiency level). The follow-up interviews took place three weeks after the computer-based test.<sup>35</sup> Nine experimental test items were selected (three for each indirect meaning category) and presented to the interview participants again. Taguchi et al. indicated that these items were selected because they “functioned well to discriminate between the high and low-scoring groups” (p. 687). However, it is not clear if the selection of items was done based on the top and bottom third groups for the overall test or if they used only answers from follow-up interview groups (eight participants total).

During the interviews, participants viewed each item one by one, and were asked again to select an answer. The researcher then asked them questions in order to gain some insight into the

---

<sup>35</sup> This was different from Taguchi (2002) for which verbal introspective reports happened right immediately after participants selected their answer. This was made possible by the relatively small number of participants used in 2002 (only eight total). In Taguchi (2009a), the verbal interviews focused on a smaller subset of the participants who had completed the computerized test and the protocol was similar to Taguchi et al. (2016).

comprehension process. Taguchi et al. provide examples of follow-up questions (p. 687):<sup>36</sup>

1) *Why did you choose this answer?*

2) *What were you thinking when you chose the answer?*

Interviews lasted about 15 minutes each and were recorded with a digital voice recorder. One researcher later transcribed them and another author checked the coding. Based on a list of strategies Taguchi had previously published (Taguchi, 2002), 72 reports were coded to determine the frequency and nature of the inferencing processes used by these low- and high-scoring participants. The following strategies were identified, presented with examples:

#### **Figure 4**

*Taguchi et al. (2016) – Categories of Inferencing Strategies and Sample Interview Excerpts*

##### **Categories of Inferencing Strategies and Sample Interview Excerpts**

Auditory cues: Identifying vocal qualities (intonation, stress, pause, tone, or speech rate) to infer the speaker's intention.

*"I thought that her tone sounded a little bit more negative than positive. So that's why I chose number 3 and not number 4."*

Visual cues: Identifying visual cues (e.g., facial expressions and gesture) to infer the speaker's intention.

*"u:h he made u:h a face at her haircut that look like dislike or disgust so I was like, nah he doesn't like it."*

Background knowledge: Referring to life experience and global knowledge to make inferences.

*"If your back hurts more after than before then it's not a good thing and you probably don't like it."*

Key word inferencing: Picking up a key word and associating it with meaning.

*"I'm pretty sure I heard sufre in the last sentence [laughs] and (2.0) I think she meant like (2.0) suffering that the trip was like a bad thing."*

Logical reasoning: Working deductively toward the intended meaning.

*"Marcela said that she didn't have enough money to even buy a jacket. So she can't go. Uhm. traveling with him to London."*

(Taguchi et al., 2016, Table 4 p. 690)

<sup>36</sup> In Taguchi (2002), examples of follow-up questions included: "Why did you choose the answer?" "Were there any hints in the conversation?" and "What were you thinking when you chose the answer?" (p. 158).



In our previous chapter, [section 3.3](#), we discussed potential issues with some of the listed strategies and the fact that they may sometimes overlap or be combined. We should note here that while Taguchi et al. state that they based their strategies on Taguchi (2002) there were some differences between the two papers. We note the addition of the “visual cues” strategy, which stems from the fact that Taguchi et al. used an audio-visual instrument when Taguchi (2002) only had a listening test. The “paralinguistic cue” strategy from Taguchi (2002) solely relied on auditory cues (which is why Taguchi et al. renamed that strategy as “auditory cues” in their paper). Two strategies present in Taguchi (2002) were not included in Taguchi et al. (2016): the adjacency pair rule and the speaker intention. In [Chapter 3](#) we discussed the possibility that the “logical reasoning” category could potentially encompass a deductive approach based on the recognition of the speaker’s intention so it is possible the two categories could be merged. While they did not specify it, we assumed that Taguchi et al. used the same data-driven approach Taguchi had used in 2002. Therefore, the fact that the adjacency pair rule strategy was not listed in Taguchi et al. (2016) may just mean that that particular strategy was not present in the L2 learners’ verbal reports. We should however keep in mind that, Taguchi (2002) links both paralinguistic cues and adjacency pair expectation to “people’s encyclopaedic entries” (p. 166) which Sperber and Wilson (1995) define as “stereotypical assumptions and expectations about frequently encountered objects and events” (p. 88). Drawing on Yus’s (2000) more detailed description of that category of contextual source, we know it to include macrosocial norms and factual information, common sense assumptions, general and prototypical expectations, and microsocial and situational expectations. We therefore suspect some possible overlap with the background knowledge category (even though Yus, 2000, clearly separates the two since co-text appears as its own contextual source in his list). Another notable difference for the logical reasoning strategy is that Taguchi et al. (2016) shorten the description used in Taguchi (2002): “recognizing literal meaning and working deductively toward the implied meaning.” They only kept the second half of that description and did not reference recognizing literal meaning. However, we suspect that they still adopted the restrictive approach of coding for logical reasoning only when that logical reasoning was based on an understanding of the literal meaning for two reasons: one, as we mentioned previously ([section 3.3](#)), any inferencing strategy could be viewed as involving some type of logical reasoning, so a broader definition would likely result in a much higher number of such strategy. Two, the sample interview excerpt Taguchi et al. provided to illustrate this strategy (c.f. [Fig. 4](#)), which they discussed in their paper, was based on the fact that the participant

“comprehended the entire utterance [target sentence], including the word [jacket], and worked deductively to arrive at the implied meaning” (p. 691). Additionally, in their discussion, Taguchi et al. also describe the use of the strategy by high-scoring participants as “processing the utterance meaning first and working deductively toward the meaning behind the utterance” (p. 692).

#### 4.4. Results and discussion

We will now take a look at Taguchi et al.’s quantitative and qualitative results. Taguchi et al.’s overarching research question was to investigate “the extent to which L2 Spanish learners’ comprehension differed across item categories in terms of accuracy and speed” (p. 688).

In order to measure accuracy, one point was granted for each correct answer to the experimental items. Therefore, a total of 12 pts could be received for each indirect meaning category (overall total of 36 pts). All incorrect answers were marked with a score of zero. Below is a summary of their comprehension accuracy scores (Taguchi et al.’s Table 2, p. 688):

**Table 2**

*Taguchi et al. (2016) – L2 Spanish Learners’ Comprehension Accuracy Scores*

	Mean	<i>SD</i>	Min.	Max
Indirect refusals	10.78	1.16	8.00	12.00
Indirect opinions	10.53	1.92	5.00	12.00
Irony	9.09	1.89	5.00	12.00
Total	30.40	4.12	22.00	35.00

Taguchi et al. indicated that participants’ overall accuracy scores fell between 75 and 90% (p. 688), which is fairly high and indicative of a good level of comprehension.

They operationalized comprehension speed as response times; the number of seconds

participants took to respond correctly<sup>37</sup> was averaged out per category. The following descriptive statistics were presented (Taguchi et al., Table 3, p. 688):

**Table 3**

*Taguchi et al. (2016) – L2 Spanish Learners' Comprehension Response Times*

	Mean	SD	Min.	Max
Indirect refusals	9.10	2.47	3.45	14.04
Indirect opinions	7.87	2.36	3.21	13.18
Irony	8.95	2.48	3.52	14.58

Since their data were not normally distributed, Taguchi et al. used non-parametric tests to analyze both scores and response times as a function of indirect meaning type.

#### 4.4.1. *Effects of indirect meaning types on comprehension*

While participants did well overall, Taguchi et al. found that the irony category seemed slightly more challenging (average of 9.09 out of 12 pts) than the other two categories (respectively 10.78 and 10.53 pts average). Indirect opinions required the least amount of time ( $M = 7.87$ ) and indirect refusals, while seemingly easier to understand (highest accuracy score), took the longest ( $M = 9.10$ ).

A significant difference in L2 comprehension was found across three indirect meaning categories using the Friedman test (chi-square = 27.66,  $p = 0.0001$ ). The Wilcoxon signed-ranks test showed no significant difference between indirect refusals and indirect opinion scores. However, irony scores were significantly lower than scores for both indirect refusals and indirect opinions (p. 689).

Taguchi et al. used Cohen's  $d$  to measure effect size and reported a large effect size ( $d = 1.01$ )

---

<sup>37</sup> In Taguchi (2005), an analysis of the response time on incorrect items was included to investigate a possible accuracy-speed trade-off. Based on previous L1 research (Taguchi references Lachman, Lachman, & Butterfield, 1979; Oomen & Postma, 2001), it was expected that increased speed would result in more errors, and that better accuracy might result in longer processing time. Taguchi's (2005) findings did not support this claim. In all but two of the other studies by Taguchi we reviewed, comprehension speed was always based on response times on correctly answered items.

in the comparison between irony and indirect refusals, and a medium-large effect size ( $d = 0.76$ ) in the comparison between irony and indirect opinions. However, effect size calculation is a point of contention among linguists right now. Plonsky and Oswald (2014) recommend adjusting benchmarks based on L2 research design, and use the following for within-subjects comparison (as opposed to between subjects):<sup>38</sup>

- small effect size when  $d$  lies between 0.6 and 0.99
- medium effect size when  $d$  lies between 1.0 and 1.39
- large size is when  $d$  is beyond 1.40.

Based on Plonsky and Oswald's (2014) scale, Taguchi et al.'s effect size would be medium for the irony-indirect refusal pair, and small for the irony-indirect opinion comparison.

While not the focus of their study, they also ran similar analyses on the native-speaker scores and reported the following results which showed a similar pattern to those of the L2 speakers: there was a significant difference across all three categories of indirect meaning (chi-square = 10.67,  $p = 0.005$ ), and irony was harder to understand than indirect refusals and indirect opinions, but there was no difference between indirect refusals and indirect opinions. A large effect size was found in the difference between both irony and indirect refusals ( $d = 1.45$ ), as well as between irony and indirect opinions ( $d = 1.13$ ).

Similar statistical tests were run on the comprehension speed data. The Friedman test revealed a significant difference in L2 response time across the three categories (chi-square = 21.44,  $p = 0.0001$ ) as well as native-speaker response time (chi-square = 11.14,  $p = 0.004$ ). The Wilcoxon Signed-ranks test showed that both L2 and native-speaker group took less time to answer indirect opinion items than indirect refusals and irony (no statistically significant difference was found between indirect refusals and irony). A medium-large effect size was found for the following two comparison in the L2 response times: indirect opinions vs. indirect refusals ( $d = 0.50$ ) and indirect opinions vs. irony ( $d = 0.45$ ). For the native speakers, a large effect size was observed in comparisons between indirect opinions and indirect refusals ( $d = 0.79$ ) and between indirect opinions and refusals ( $d = 0.83$ ).

In their discussion section, Taguchi et al. argued that the fact that irony was more difficult to comprehend may relate to its “greatest distance between the utterance meaning and the intention

---

<sup>38</sup> due to the fact that effects are usually larger for within-group contrasts

behind the utterance” (p. 692). These results were in line with findings published by other authors (citing Bouton, 1992, 1994b, Yamanaka, 2003, and Shively et al., 2008). However, irony items required as much time to comprehend as indirect refusals (both were categorized as conventional). Taguchi et al. state that irony requires the listener to attend to available contextual sources (visual cues, background knowledge, auditory cues, etc.) to recognize the mismatch between the propositional meaning and intended meaning, as well as the attitudes and emotion of the speaker: “without processing these multiple cues, i.e., linguistic, nonverbal, and contextual, learners might wrongly jump into the literal meaning” (p. 693). Yet, multiple times, Taguchi has said that this multi-level processing increases the overall processing effort and is often seen in longer response times. This means that there is an apparent mismatch in the results reported by Taguchi et al.: high levels of accurate irony comprehension (suggesting the processing of multiple cues) without concomitant longer response times. One possible explanation of the shorter-than-expected response times that Taguchi et al. put forward is the salience of the audiovisual cues, which made it possible for learners to process the irony items quickly, especially “when they understood the convention of this rhetorical device” (p. 693).

Now we turn our attention to the comprehension of indirect opinions. Taguchi et al. classified them as nonconventional, and the assumption was that they would be harder to understand and would require more time than the conventional implicatures because their indirect mechanism was more idiosyncratic, and therefore people would need to process more contextual cues to reach the correct interpretation. Taguchi et al. however found that indirect opinions had the fastest comprehension speed. After analysis of their learners’ strategies (which we will discuss shortly), Taguchi et al. hypothesize that the reason why learners took less time to answer indirect opinions was that the learners used visual cues to process indirect opinions at a much higher rate than was the case when processing other types of indirect meaning. Taguchi et al. explain that “use of visual cues is more economical because visual information directly maps onto meaning. As a result, listeners can bypass the bottom-up processing of an utterance, which results in shorter response times” (p. 691). Given that the indirect opinions in Taguchi et al. were majoritarily negative opinions, it is easy to imagine that the speakers would have expressed their discontent, dislike or disapproval with corresponding gestures or facial expressions that may have been easily recognizable (especially if combined with a conventional turn of conversation as we have discussed previously, which would have primed the L2 learners in expecting some kind of positive or negative opinions).

With regard to the impact of conventionality, a comparison of indirect refusals (conventional items) and indirect opinions (nonconventional implicatures) did not support Taguchi’s previous findings (2005, 2009b, 2011) in which she had observed that “indirect refusals were easier and faster to comprehend than indirect opinions” (Taguchi et al., 2016, p. 693). A possible explanation provided by Taguchi et al. was the relatively high level of proficiency of the L2 participants used in the 2016 study: indeed, they were able to comprehend both indirect refusals and indirect opinions had a very high rate (88 to 90%) which was much higher than the accuracy rates obtained in Taguchi’s 2011 study on L2 English. They hypothesized that using participants at a lower level of proficiency (such as beginner students) may have yielded results more similar to those seen in Taguchi’s (2008a) study (using L2 Japanese learners).

#### 4.4.2. *Posthoc Analyses of Retrospective Verbal Interviews*

As mentioned previously, Taguchi et al. collected a total of 72 verbal reports (from eight learners: four low-scoring participants and four high-scoring participants who were asked to verbalize their inferential process while reviewing a total of nine items—three for each category of indirect meaning). In total, they identified a total of 100 strategies. Per Taguchi (2002), “a discrete strategy was counted only once per item, even if it appeared repeatedly within each item” (p. 161). Their data included 55 strategies used by the high-scoring group and 45 strategies used by the low-scoring group. The table below presents Taguchi et al.’s distribution of strategies based on performance (from Taguchi et al, Table 5, p. 690):

**Table 4**

*Taguchi et al. (2016) – Frequency of Inferencing Strategies by Participant Group*

	High-scoring learners (n=4)	Low-scoring learners (n=4)
Auditory cues	7 (12.7%)	5 (11.1%)
Visual cues	5 (9.1%)	5 (11.1%)
Background knowledge	4 (7.3%)	0
Key word inferencing	7 (12.7%)	17 (37.8%)
Logical reasoning	32 (58.2%)	17 (37.8%)
Total	55	45

Taguchi et al. reported that high-scoring participants “used a wider range of strategies and used strategies more frequently than low-scoring students” (p. 689). High-scoring participants

reached a high level of accuracy (97%) while low-scoring students only reached the correct interpretation 70% of the time. Taguchi et al. did not report on possible discrepancies in the answers the interview participants provided as compared to their initial answers on the computerized test.

Both groups of participants used paralinguistic cues and reported noticing specific facial expressions, gestures, etc. There were however some differences between low- and high-scoring students. First, we can see in [Table 4](#) that low-scoring students used key word inferencing more often than high-scoring students. On the other hand, high-scoring participants used logical reasoning more than any other strategies and also more than the low-scoring students (a result that departs from Taguchi's [2002] study in which she found that both high proficiency and low proficiency participants used logical reasoning at a similar rate). We explained in [Chapter 3 \(section 3.3\)](#) why the logical reasoning category may be problematic, both because we believe that any inferential process will require some level of logical reasoning, but also because identification of this strategy is made difficult by the ability to verbalize such process (or lack of verbalization). As previously mentioned ([sub-section 4.3.3](#)), we suspect Taguchi et al. only coded for logical reasoning when the participants first recognized the literal meaning of the target utterance before logically inferring the intended meaning. This would explain why Taguchi et al. did not find a very high frequency of use of logical reasoning strategy in their lower-scoring participant group, since they often only heard keywords as opposed to the entire utterance or dialogue.

Surprisingly, none of Taguchi et al.'s low-scoring learners reported relying on background knowledge to comprehend indirect meaning. We may remember that, in a previous study, Taguchi (2002) had found that low proficiency participants used this category of strategies at a higher rate than the high proficiency participants. Taguchi et al. did not provide an explanation for the lower use of this strategy (overall but particularly for low-scoring participants), or for the discrepancy with Taguchi's (2002) findings, though they seem to allude to the fact that background knowledge is often used for more complex implicatures such as irony. The low-scoring participants' inability to comprehend irony may therefore be linked to their lack of use of this strategy.

In their paper, Taguchi et al. also provide excerpts of the interview transcript showing the use of specific strategies to illustrate the inferential process and success of that strategy. In one

particular excerpt (excerpt 2, pp. 690-691), Taguchi et al. showed how “one learner relied on the wrong key word, which led to a wrong inference” (p. 690). Of course, this predominance of key word inferencing strategies in low-scoring students is likely linked to their lower proficiency. As we know, lower proficiency students are more likely to focus on linguistic meaning, particularly lexical items they may be familiar with: their inability to comprehend the full dialogue leads to their reliance on the few words they do understand, and the increased attention to linguistic decoding may make it more challenging for them to attend to other cues. Further evidence of this process is shown in Taguchi et al.’s third excerpt in which a participant “explicitly said that she was tuning to a key word, specifically that she was able to catch the key word (*good*), but she was not confident, ending up with the wrong selection” (p. 691). In contrast, high-scoring students who were able to understand the full dialogue, were able to use their deductive skills to reach the intended meaning interpretation.

In addition to looking at differences in strategy use among high and low-scoring participants, Taguchi et al. also wanted to investigate potential differences among the implicature types. The table below shows the distribution of strategies depending on the type of indirect meaning (from Taguchi et al. 2016, Table 6, p. 691):

**Table 5**

*Taguchi et al. (2016) – Frequency of Inferencing Strategies by Item Type*

	Indirect refusals	Indirect opinions	Irony
Auditory cues	1 (3.2%)	3 (8.8%)	8 (22.9%)
Visual cues	1 (3.2%)	6 (17.6%)	4 (11.4%)
Background knowledge	1 (3.2%)	2 (5.9%)	1 (2.9%)
Key word inferencing	8 (25.8%)	5 (14.7%)	11 (31.4%)
Logical reasoning	20 (64.5%)	18 (52.9%)	11 (31.4%)
Total	31	34	35

As seen in the above table, auditory cues were used more for irony than for the other two types of implicatures. Several learners reported picking up on the ironic or sarcastic tone of the speaker, which helped them interpret the intended meaning. Taguchi et al. also note that key word inferencing and logical reasoning were used at a similar rate for the interpretation of irony by the learners, “suggesting that they used multiple strategies simultaneously to comprehend



irony” (p. 691). According to Taguchi et al., the need to attend to multiple cues was necessary to interpret irony since it “represents the greatest distance between the two levels of meaning (sense and force)” (p. 692), and in turns, that higher processing demand explained why accuracy was lower for irony.

In contrast, indirect refusals and indirect opinions were often interpreted primarily with the use of a single strategy: logical reasoning. Taguchi et al. then explain that “because these two item types were easier to comprehend than irony, as indicated by higher accuracy scores, it makes sense that comprehension was relatively straightforward with no need for using multiple strategies” (p. 691).

Taguchi et al. also note that visual cues were used more often for indirect opinions and concluded that access to these visual cues factored in the faster response times their participants had for that category of indirect meaning (a finding that was different from Taguchi’s previous studies, but Taguchi et al. also pointed out that her prior experiments did not rely on videos, only listening tests). According to Taguchi et al, visual cues are quickly accessible and the strategy that relies on them will be faster than other types of strategies that involve more complex processing (like logical reasoning): “Different from logical reasoning that requires processing of individual utterances, use of visual cues is more economical because visual information directly maps onto meaning” (p. 691). An example of this was provided by Taguchi et al. who reported on a learner who did not process the target utterance of an indirect opinion but simply “picked up on the speaker’s facial expression and immediately used it as evidence to infer the underlying intention” (p. 692).

Summarizing their findings regarding the variety of inferencing strategies used by their participants, Taguchi et al. reiterate the importance of context in the search for relevance when dealing with indirect meaning and the possibility and ability of L2 learners to access and use these contextual sources. The prevalence of paralinguistic cues in the interview data also support the use of a multimodal instrument that presents audio-visual input for a more realistic processing of the indirect meaning. At the end of their paper, Taguchi et al. also offer some pedagogical suggestions. Given the noted difficulties learners encountered with irony items, explicit classroom instruction on irony would be beneficial. Specifically, instructors may explain the conventional pattern of irony (meaning reversal/replacement) and draw attention to paralinguistic cues that often accompany this type of implicature. Use of audio-visual material

to illustrate this phenomenon is of course recommended, as well as a series of activities around cross-cultural and cross-linguistic comparisons and discussions about the intent and communicative effects of the implicature. Encouraging students (in particularly low-proficiency learners who tend to rely on key word inferencing too much) to use a variety of strategies (such as attention to paralinguistic cues) and focus on the overall context rather than “word-level comprehension” (p. 695) should “help learners make use of a range of tactics that support intelligent guessing of indirect meaning” (p. 695).

#### 4.5. Limitations and potential for replication and extension

Throughout our review of Taguchi et al. (2013), we have discussed potential methodological issues. These problematic design elements will have to be addressed in the scope of our replication (and will be discussed in [Chapter 5](#) on our instrument development, as well our quantitative methods sections [6.1](#) and [7.1](#), and [Chapter 8](#) on the qualitative methods). Indeed, part of the reason why we wanted to replicate Taguchi et al.’s study was to see if we could corroborate some of their findings and generalize them to another language. This means that part of our dissertation attempted to carry out a close replication of Taguchi et al. (2016). Yet, as we will discuss shortly, we saw an opportunity to exploit our collected data and conduct additional analyses that could further our understanding of L2 indirect meaning comprehension and potentially remedy to some of Taguchi et al.’s drawbacks.

Taguchi et al. recognize some of the limitations of their study: first, the fact that their sample size was small implies that replication or larger sample size studies are needed to be able to generalize their findings. They also explain that the dialogues used in their study had to be adapted for the lower-proficiency learners and therefore do not reflect completely naturalistic discourse. In their recommendations for future research, they suggest using a wider range of proficiency levels (since their participants had a fairly high level of proficiency which Taguchi et al. thought was responsible for the lack of conventionality effect) which need to be assessed with reliable metrics,<sup>39</sup> and further investigate the impact of conventionality in various forms. The role of visual cues to facilitate comprehension should be studied in the same manner as linguistic conventions. Finally, “the interaction among proficiency, conventionality and

---

<sup>39</sup> like the ACTFL Oral Proficiency Interviews

comprehension of indirectness [...] would support a deeper understanding of established general inferential mechanisms across languages” (p. 696).

There were a few other points in Taguchi et al.’s study that we felt could be explored further. First Taguchi et al. reported that their native speaker group exhibited similar patterns of comprehension compared to the L2 learner group—in particular, both groups seemed to have more difficulties with irony comprehension. If that is the case, that could indicate that the issue with indirect comprehension is not specific to L2 learning and processing, but has more to do with comprehending indirect meaning in general. One way to test this hypothesis would be to compare L2 learners’ comprehension of indirect meaning in both their L2 and their L1 (using similar tasks) to see if the difficulties in comprehension were directly related to the foreign language. Of course, a close replication of Taguchi et al. (2016) would not allow us to collect data on the learners’ L1 comprehension, but the expectation was that our findings with regarding to native (pilot study) and non-native performance would be a way to generalize (or disapprove) Taguchi et al.’s findings regarding L1 comprehension.

Second, no analysis of the literal meaning items was provided in Taguchi et al.’s study (and Taguchi used literal items in many of her other studies but never investigated possible differences in the understanding of literal meaning as opposed to the indirect meaning). We may wonder whether Taguchi et al.’s participants would have had issues understanding literal meaning as well (for example a direct refusal or direct opinion). Contrasting an ironic statement with the identical statement meant to be interpreted literally may have shed some light on the L2 processing of irony and corroborated various claims regarding literal meaning processing. This was one aspect that we thought our close replication could expand upon Taguchi et al. (2016). Since Taguchi et al. included literal meaning items as fillers, we could still follow the same study design and protocol, but include the literal meaning items in our analysis of the collected data. Without comparing comprehension of literal items with the comprehension of non-literal, it would be hard to make claims as to the difficulties of indirect meaning comprehension, since the issues could be linked to general comprehension.

As we mentioned in [section 4.3](#), Taguchi et al.’s participants in the retrospective verbal interviews were selected based on their accuracy scores on the multiple-choice questionnaire. We do not know whether low-scoring participants were lower-proficiency students and whether high-scoring participants had reached a higher level of overall proficiency because the

proficiency information Taguchi et al. collected was never used in the data analysis—only to provide descriptions of the participants. Indeed, despite the fact that Taguchi et al. used students enrolled in intermediate and advanced Spanish classes at one American university, proficiency or institutional-level groups were never constituted to reflect on the impact of general L2 proficiency, and students' results on the NCSSFL-ACTFL Can-do self-assessment tool were never computed to differentiate between low and high-proficiency levels.

Finally, while the identification of inferential strategies used by L2 learners was invaluable and provided useful insights into the cognitive processing of indirect meaning in L2, the analysis of frequency and distribution for each type of implicatures fell short. According to us, two things were missing. First, it would have been useful to know whether differences between low- and high-scoring students were observed in their processing of the different kind of implicatures (not just overall). In other words, when looking at strategies for each indirect meaning type, Taguchi et al. did not distinguish between high- and low-scoring participants. Second, Taguchi et al. did not present information about the strategy success rate: which (categories of) strategies and how many were successful in reaching the correct interpretation? We propose to attempt to tackle these two points in our replication.

## Part I - Conclusion

In this first part of our dissertation, we presented an overview of replication, its importance and potential challenges ([Chapter 1](#)). In [Chapter 2](#), we provided a theoretical framework for indirect meaning: what implicature means, the different types of indirect meaning and how people process implicatures. In [Chapter 3](#), we reviewed existing research in L2 comprehension (and to a lesser extent production) of indirect meaning and how factors such as proficiency, types of implicatures, degree of indirectness or conventionality, and access to contextual cues can affect comprehension. We also looked into studies that investigated inferential strategies L2 learners used to understand indirect meaning, and presented a list of these strategies and how they relate to contextual cues. One aspect which we wanted to emphasize is the interaction of different factors: for example, comprehension of a specific type of implicatures may be more challenging for low-proficiency learners who may then select different types of strategies to compensate for the relative opacity of the implicatures. Finally, in [Chapter 4](#), our analysis of Taguchi et al.'s (2016) findings showed some potential methodological issues, and opened the door to a further study of indirect refusals, indirect opinions and irony in L2 comprehension.

Following Taguchi et al.'s (2016) recommendation to conduct “more research on languages other than English” (p. 680), we set out to replicate Taguchi et al.'s study to “confirm cross-linguistic generalizability” (p. 680) with a group of L2 French learners. The overall purpose of this replication, as stated by Taguchi et al. (2016) is to “test theoretical claims about inferencing mechanisms and to understand the construct of indirectness comprehension in an L2” (p. 680) other than Spanish (which was the target language in Taguchi et al., 2016, as well as Koike, 1996, and Shively et al., 2008), Japanese (Taguchi, 2008a; Taguchi, 2009a), Chinese (Taguchi et al., 2013) or English (focus of many studies including Bouton, 1992, 1994a, 1999; Taguchi, 2002, 2005, 2008c, 2011, Takahashi & Roitblat, 1994; Cook & Liddicoat, 2002; Yamanaka, 2003; Garcia, 2004).

Our experimental study seeks to answer the following research questions:

- 1) To what extent are L2 French learners able to understand indirect meaning?**
- 2) How do different types of indirect meaning impact L2 French learners' comprehension accuracy and speed?**

### **3) What inferential strategies do L2 French learners use and how successful are they?**

[Part II](#) of this dissertation will try to address our first two research questions using quantitative data collected via an instrument and protocol similar to that of Taguchi et al. (2016). Our findings with regard to indirect meaning comprehension (including a comparison of L2 learners and native speakers) will be compared to Taguchi et al.'s findings to determine if there were differences, or if their findings were reproducible and can therefore be generalized. We also propose to extend their study by including results for the literal meaning items (fillers) in our statistical analysis in order to contrast comprehension of indirect meaning and comprehension of more direct meaning.

[Part III](#) will focus on our third research question and use a more qualitative approach to analyze transcripts from retrospective interviews. First, we will include a review of strategies that were present in Taguchi et al. (2016), replicating their coding protocols. We will analyze the frequency and use of strategies by each participant group as well as a function of indirect meaning types. Going beyond Taguchi et al.'s analysis of strategy selection for each category of indirect meaning, we will also look at the distribution of the combined factor (group and meaning type) and include a discussion on strategy success.

## **PART II**

# **EXPERIMENTAL STUDY**

## Introduction

Taguchi et al.'s (2016) study adopted a mixed method paradigm to answer their research questions. Part II of our dissertation will be devoted the quantitative aspect of our replication of their study.

[Chapter 5](#) will be dedicated to our instrument development: we will describe how we developed the multimedia test (section [5.1](#)), based on Taguchi et al.'s (2016) study, as well as the online survey used to collect our data (section [5.2](#)). We will also address some of the challenges we encountered when creating our instrument.

[Chapter 6](#) will focus on the pilot study with a group of native speakers. We will first discuss our methodology (section [6.1](#)), and then present results for the native speakers' comprehension accuracy and speed so as to compare them with Taguchi et al.'s findings (section [6.2](#)). In [section 6.3](#) we will explain how the results of this pilot study were utilized to refine our instrument.

Finally, in [Chapter 7](#) we will attempt to answer our first two research questions:

- 1) To what extent are L2 French learners able to understand indirect meaning?**
- 2) How do different types of indirect meaning impact L2 French learners' comprehension accuracy and speed?**

To that effect we will present a detailed account of the methodology we follow for the L2 learner experiment (section [7.1](#)). We will present our results in terms of accuracy and speed (section [7.2](#)). In order determine whether comprehension difficulties stem from foreign language processing or if indirect comprehension issues are experienced by both native speakers and L2 learners alike, we will reprocess our native speaker data to only include tasks that were used in the L2 learner experiment so as to compare the two groups (section [7.3](#)). Finally, in [section 7.4](#) we will discuss our results as well as potential differences in findings with Taguchi et al.



## Chapter 5 - Instrument development in the context of a replication study

In this chapter, we will present the methodology we used to develop our instrument. We will first detail the various steps we followed to develop the multimedia test (including the translation of the non-literal items, issues we encountered during this process, the feedback we received and modifications we made, and the creation of the fillers, and finally, our process for filming our video clips). We will then turn our attention to the development of the data-collection instrument: starting with the multiple-choice questionnaire (which also involved a translation of Taguchi et al.'s questions), and the instructions we provided participants. Finally, we will explain the challenges we faced when designing our online platform (along with the solutions we adopted).

Since our project aims to replicate and expand upon Taguchi et al.'s study, it was important for us to try to stay as close as possible to the original design of the study that Taguchi et al. published. We contacted Taguchi and asked if she would be willing to share the dialogues and tasks she and her team designed. She graciously sent us the “final version” of the pilot test she used. The file included the experimental dialogues as they were initially designed, as well as the associated multiple-choice questions (in both Spanish and English). A sample of these test items was also published in their paper. The document we received presented 48 experimental items (16 for each of the indirect speech categories). It did not include the fillers, neither did it indicate which items were later excluded for the final version of the survey used in study. Taguchi et al.'s pilot study had 48 experimental items and 10 fillers.

The final version of our items (48 total, including the experimental items that were later discarded for the L2 learner test) both in English and French can be found in [Appendix C](#).

### 5.1. Multimedia Listening Test Development

The following sub-sections will give an account of the multimedia test design process, from translating Taguchi et al.'s dialogues, to creating fillers, seeking feedback from native speakers and producing the video clips.

### 5.1.1. *Non-literal dialogues*

In order to create dialogues for the multimedia test, we considered two different approaches. The first one was to follow the same process Taguchi et al. undertook, and use personal notes from field observations, scenes from movies or TV shows, or conversations found in existing oral corpora, to come up with original dialogues. However, given that we had specific criteria to follow to match Taguchi's test items for the replication (i.e., the type of indirect speech used, similarity of situations, length of dialogues, etc.), we decided to use the dialogues that Taguchi and her team created as the basis for our work. As will be shown later, this decision was later supported by native speakers (from whom we sought feedback) as well as our pilot study.

We began by translating the original dialogues and checking for appropriateness. We used both the Spanish and English versions provided by Taguchi, but had to make some changes to ensure the dialogues were culturally relevant in a French context. Just like Taguchi et al., we made sure to use vocabulary and grammar that would be accessible to lower-level learners. That is not to say that we assumed all learners at that level would comprehend every word of the dialogues, but in essence, we tried to limit the vocabulary and structures to those that were present in our elementary and intermediate-level textbooks,<sup>40</sup> paying particular attention to the last part of the dialogues in which the indirect utterances were present. Since the focus of this project was to explore learners' comprehension of non-literal language, we needed to try to eliminate, as best as we could, the possibility that learners may not understand the indirect speech due to their inability to understand individual lexical and grammatical items. There is obviously very little chance someone could understand the intent of an ironic comment, if they were not able to understand the words composing that ironic sentence.

For the sake of consistency and comparability, we named each dialogue using Taguchi et al.'s labelling system: IRF for the indirect refusals, IRO for Irony and OPI for indirect opinions. The three letter prefixes were then followed by a number (from 1 to 16 for the experimental items) and the French short title corresponding to the main topic of the dialogues.

---

<sup>40</sup> Manley, Joan H. *Horizons*. 7<sup>th</sup> ed., 2019. Cengage Pub.

Muyskens, Judith., Harlow, Linda., Vialet, Michèle., & Brière, Jean-françois. *Bravo!* 8<sup>th</sup> ed., 2015. Heinle & Heinle Pub.

### 5.1.2. Translation choices and modifications

As mentioned before, when creating the French dialogues, we tried to stay as close to the original conversations as possible. However, at times, translation was challenging and we had to make some changes. To confirm plausibility of the conversations and accuracy of the vocabulary and grammar, all our translations were verified by four native speakers from whom we sought feedback. We will present that feedback in a later section. Here is a list of elements we had to take into consideration to decide whether to make some general changes to the Spanish version created by Taguchi et al.:

- *Names of people:* When asked via email, Taguchi indicated that the same two actors were employed for all the dialogues: one male and one female speaker. The original study however used different female and male names in the dialogues. All names were Spanish-sounding, so in a true replication fashion, we kept a similar pattern but used French-sounding names. For example, OPI7 in Taguchi et al.'s study features a male speaker named *Germán* and a female speaker named *Sofía*. Our corresponding French dialogue introduces two characters named *Gauthier* and *Sophia*.
- *Characters' gender:* Each of our dialogues included one male and one female (in their early thirties) character. Two of Taguchi's dialogues (OPI1 and IRO1) were centered around two male characters (though the actors were still supposed to be male and female): we decided to switch to one man and one woman to preserve the system in place. Since the document Taguchi had sent also included items that were later rejected, there was no way to know whether the dialogues with two men may have been the ones that they decided not to use in the final study. An example of this is dialogue IRO1: the dialogue takes place between a father and his son in Taguchi et al.'s version, whereas ours revolves around the interaction between a father (male actor) and his daughter (female actress).
- *Characters' age:* Some of Taguchi et al.'s dialogues seemed to assume a relatively small age difference between characters (e.g., two friends, two classmates, a husband and a wife). However, two conversations implied a more significant age gap (such as IRO1, featuring a parent and a child<sup>41</sup> or IRF6 between a student and her advisor/professor). Since Taguchi had indicated that her team only used two actors for the entirety of the videos, we decided not to make any changes to this aspect of our study.

---

<sup>41</sup> In our dialogues Fill3 also involves a conversation between a mother and her son

- *Names of places:* Again, here since the dialogues were transposed in a French setting, we had to change the name of restaurants to ones that sounded French, or that were culturally more appropriate. For example, the sandwich place called the “*Museo del Jamón*” (or the “Proscuitto Museum” as translated by Taguchi et al.) became “*au Bon Pain*” in our French version.
- *Cultural changes:* while we tried to be as faithful to the original dialogues as possible, at times, using a direct translation was deemed culturally inappropriate for our French context. Spanish and French cultures have a lot of similarities and certainly, when it came to indirect meaning, it seemed that the items designed by Taguchi et al. still worked once transposed to French. However, there were instances when specific situations or elements of the dialogues did not make as much sense in a French context. Here are some examples of modifications we made on that basis, either because the direct translation seemed culturally irrelevant, or because cultural practices are different between France and Spanish-speaking countries:
  - For IRF4, we used “*modules*” to refer to learning units/modules in both our French and English versions because we wanted to avoid using the word “unit” used in Taguchi’s test so as to eliminate any confusion with term “*Unités d’Enseignement*)” that is commonly used in French universities to refer to course categories.
  - Dialogue IRF6 introduces a student and her advisor in the original study. However, given that the student is inquiring about an assignment she had to complete for her class, we thought it made more sense to replace the “advisor” by a professor. The Spanish version also used first name basis in the dialogue, which we thought more inappropriate and unlikely to happen in a French context where undergraduate students usually refer to their professors by their last name. Therefore, we replaced the informal address and first name “Martin” by a more formal “*Monsieur Bodnar*” and “*vous*” pronoun address.
  - Quite a few edits were made in dialogue IRF9. First, the two characters in the original dialogue are talking about going out to eat some cake at a restaurant or bakery. While it is common practice in the US to go out for dessert, it is quite unusual for French people to go to a restaurant to eat cake. Therefore, we decided to set the scene a little bit differently (while still preserving the intended meaning of the final sentence). Instead of going out for cake, our characters are making plans to try out a new all-you-can-eat buffet at the local “*China King*” (which is

a common Asian restaurant name in France). The times were switched to a more appropriate 24-hour system of 19h and 19h15, which also accounted for the fact that French people may eat later. Finally, the Spanish version of this dialogue refers to the female speaker buying movie tickets in advance. This is not very customary in France, as most people usually buy their tickets when they get to the movie theater. Therefore, we chose to have our female speaker say that she had bought tickets to a concert (which French people do usually buy in advance). We chose “Drake” because a search on a popular French website, where one can buy tickets, showed that he was a known artist in France, and we thought our American participants would also know who he is, which means comprehension would not be hindered by this reference.

- To make IRF14 culturally relevant while making sure learners would still understand the vocabulary used, we replaced the mention of “pretzel” by “ice cream” (“*une glace*”). Indeed, this dialogue is set in a movie theater where the characters are discussing various foods they could purchase before the movie starts. Pretzels are not usually sold in movie theaters in France. Options are normally pretty limited there: besides popcorn (which was already used in the dialogue) and candy, ice cream seemed like a reasonable replacement given the content of the dialogue.
- In the original dialogue IRO8, two students are discussing their college admission (“*aplicaciones a universidades*”). One student asks the other whether she has already sent out her applications, to which she replies that she has, and inquires as to whether her friend has heard back from the universities he applied to. The college application experience is however very different in France, where undergraduate programs do not usually have such a selection process in place. For that reason, we switched the reference to the “*grandes écoles*,” which are more prestigious Ivy-league-type schools in France. Students interested in attending these schools usually have to apply in advance, and admission is competitive and not guaranteed. While there may be some concerns as to whether the American participants in our study may or may not be familiar with this concept, we anticipated that even intermediate-level students would still be able to understand “*écoles*” and equate the reference to schools or colleges.
- Both Spanish and English versions of OPI12 in Taguchi et al.’s test mention a genre of music called “chill out,” which according to Partridge and Moberg

(2017) refers to a style of “electronic ambient music” that originated in the US in the mid to late 1980s. Having never heard of this genre before, and after informally polling a dozen French native speakers who were also unfamiliar with it, we thought it best to replace it with a different genre of music altogether. We used “*la musique techno*” as we anticipated that it would be easily understood by American learners and still relevant within a French context.

- Another type of modifications we made aimed at increasing comprehensibility. Knowing that the dialogues had to be accessible to intermediate-level students, a direct translation was sometimes problematic. Here are examples of such lexical or grammatical changes:
  - In the original dialogue IRF5, two co-workers are discussing their lunch options. The male speaker Ramon says: “*Ah, yo estaba pensando en comer al aire libre en la nueva hamburguesería cerca de la playa*” (translated by Taguchi et al. as “Oh. I was thinking about eating at the new burger place in the open air next to the beach.”). In our version, we eliminated the mention of the “burger” and replaced it with “*nouveau restaurant avec une terrasse sur la plage*” (“new restaurant with a patio right on the beach.”). There was no easy way to translate “*hamburgueseria*” in French and we did not think that the reference to the burger was needed for this particular dialogue. However, we added the word “*terrasse*” because it seemed more culturally appropriate. In terms of length, the French dialogue (41 words) remained close to the original Spanish version (42 words).
  - At the end of dialogue IRF8, the female speaker in the Spanish version says that she has plans to “*ir a pilates con una compañera de clase.*” While pilates is not an unfamiliar exercise in the US or in France, we worried that the French pronunciation of the word may not be easily accessible to American learners. Additionally, there is no way to know whether all the participants would actually know what it refers to, even if they heard the word properly. Since we did not want this lexical item to be the source of misunderstanding, we decided to replace it with “tennis” since the word works in both French and English.
  - The two characters (two roommates) in Dialogue IRO3 discuss the amount of work they had to do that day. The male speaker inquires about a presentation the female speaker had to do that day. The female speaker then asks her roommate about his day. His comment that he will “*have to stay up late*” seemed hard to understand in Taguchi et al.’s version because it did not provide any

explanations as to why he would need to stay up. Since the beginning of the conversation referenced the idea of a heavy workload, we decided to clarify the situation with a little bit more information so as to avoid any confusion. We added the following to our French dialogue: “*j’ai beaucoup de devoirs donc je vais devoir me coucher tard ce soir!*” (“I have a lot of homework so I’ll have to stay up late.”). Despite the addition, our dialogue was still well within the average word count. Another issue was the direct translation of “to stay up late” in French. However, our translation was actually closer to the Spanish “*hoy me tocará acostarme muy tarde*” dialogue created by Taguchi et al. Using both the Spanish and English versions of their test was therefore quite helpful when making translation decisions. Sometimes, we had to use modulations to convey the same meaning with expressions that we anticipated would be more accessible to the lower-level learners.

- In IRO5, we replaced the question “*¿Qué lugares vas a visitar?*” (“what places will you visit”) with a more accessible question structure : “*Où est-ce que tu vas aller ?* » (“where are you going to go?”), avoiding the use of « *endroits* » which, in our experience, has a tendency to cause problems for intermediate-level learners.
  - The expression “my car broke down” in IRO13 was syntactically reversed and became “*J’ai un problème avec ma voiture*” in our French version. While our intermediate textbook presents the expression “*être en panne*,” we decided that using an expression with the cognate “problem” combined with the more accessible and frequently used lexical item “*voiture*” would better ensure learners’ ability to understand the sentence on a linguistic level.
  - The Spanish equivalent to “*baignoire*” (bathtub) was used in OPI13. Because of the low frequency of this word in intermediate textbook as well as its more complex pronunciation, we chose to replace the term by a different household item that could fit the conversation (the original dialogue mentions that the bathtub is broken and needs to be fixed). We decided to use “*lave-vaisselle*” (dishwasher) because we anticipated that at the time the intermediate-level learners would be taking this test, they would have just completed a module that uses this word frequently.
- A few other items were tagged as a potential source of concerns during the translation process (either due to potential level of difficulty for lower-level learners or for possible

awkwardness in the translation), but we chose to leave them as such, as there were no easy translation alternatives. In some instances, the more challenging elements appeared at the beginning of the dialogues and we knew they should not interfere with the learners' ability to understand the indirect meaning which came at the end of the dialogues. Additionally, the use of a multimodal test meant that participants would also have visual input to help with comprehension. For example, in IRO2, the female speaker tells her roommate: “*Est-ce que tu es prêt? Vite. Dépêche-toi.* » (“Are you ready? Come on, hurry up!”). Both “*prêt*” and the imperative form of “*se dépêcher*” in French may be hard for intermediate learners to grasp. However, since they are both used at the beginning of the dialogue, we felt that they should have a limited impact on the participants' ability to understand of the rest of the conversation. A similar situation was found in IRO12 (use of “*cauchemars*” at the very beginning of the conversation) and a few other dialogues.

The expression “*gonflé*” used in reference to one of the characters' ankle at the end of dialogue IRO10 was much more problematic as it comes right before the non-literal comment, and intermediate learners are very unlikely to have encountered the word before. However, based on the video (the man comes back from jogging limping), the provided written context (available above the video that reads: “They are talking about Emmanuel having sprained his ankle while jogging.”) and the fact that the translation of “*gonflé*” is used in three of the multiple-choice questions (“Manuel noticed that his foot was very swollen.” “Manuel did not know that his foot was very swollen.” “Camille did not notice that Manuel's foot was very swollen”), it seemed that learners should be able to infer its meaning. A similar reasoning was used in favor of keeping as close a translation as possible for the expressions “*coupe de cheveux*” and “*coiffeur*” in OPI6, “*bas*” in IRO1, “*allume*” in IRO4, and “*craie*” in OPI15.

- Through the development phase, we tried our best to be consistent and stay as close to the original Spanish conversations as possible. However, there was one dialogue that really puzzled us in its format and that we decided to modify so that it would fit the original design better. For the sake of simplicity, here is a side-by-side comparison of the two English versions (Taguchi et al.'s and mine). The Spanish and French dialogues were very close translations to their corresponding English counterparts. This also represents a good example of the overall translation process we underwent. We will draw attention to various elements of the dialogues that have been highlighted (bold, italics or underlined) below.



**Figure 5**

*IRO4 – task comparison between Taguchi et al. (2016) and our version*

<b><i>Taguchi et al. IRO4télévision</i></b>	<b><i>Our version IRO4télévision</i></b>
<p>This is a conversation between two roommates, Abel (male speaker) and Elsa (female speaker). They are talking about watching a TV show together.</p> <p><i>Abel:</i> Elsa, <b>do you have plans</b> for tonight?  <i>Elsa:</i> No, why?  <i>Abel:</i> Because there’s a new TV show that I’d like to watch. Do you want <b>to join me</b>?  <i>Elsa:</i> You know me. <u>I find TV really educational.</u> <i>Every time someone switches it on, I go to my room and read.</i></p> <ol style="list-style-type: none"> <li>1. Elsa thinks that TV is not educational. (CORRECT)</li> <li>2. Elsa thinks people can learn a lot from watching TV. (Principle 1)</li> <li>3. Elsa watches educational programs on TV every single night. (Principle 2)</li> <li>4. Elsa and Abel are going to watch a new TV show together tonight. (Principle 3)</li> </ol>	<p>This is a conversation between two roommates, Abdel (male speaker) and Elsa (female speaker). They are talking about watching a TV show together.</p> <p><i>Abdel:</i> Elsa, <b>what are you going to do</b> tonight?  <i>Elsa:</i> Nothing, why?  <i>Abdel:</i> There’s a new TV show that I’d like to watch. Do you want to <b>watch it with me</b>?  <i>Elsa:</i> You know me, <i>every time someone switches the TV on, I go to my room and read.</i> <u>I find TV really educational.</u></p> <ol style="list-style-type: none"> <li>1. Elsa thinks that TV is not educational. (CORRECT)</li> <li>2. Elsa thinks people can learn a lot from watching TV. (Principle 1)</li> <li>3. Elsa watches educational programs on TV every single night. (Principle 2)</li> <li>4. Elsa and Abdel are going to watch a new TV show together tonight. (Principle 3)</li> </ol>

The first thing one can notice in our translation is that the language was further simplified in some areas (either on a grammatical or lexical level as exemplified by the phrases in bold). For example, we thought that the French translation of “have plans” (“*est-ce que tu as des projets*”) would be much harder to understand than the perfectly acceptable, yet much more accessible, “**what are you doing to do tonight?**” (“*qu’est-ce que tu vas faire ce soir?*”). As discussed in [sub-section 4.14](#), there were a few issues with this particular dialogue, one of them being that, based on the original design, the last part of the experimental dialogues was supposed to contain the indirect meaning. However, the target ironic comment in Taguchi’s version (as reflected in multiple-choice question answer #1—the CORRECT answer— “Elsa thinks that TV is not educational”) is not the final sentence. Indeed, the non-literal part of that final statement is actually “I find TV really educational” (when Elsa actually thinks otherwise) which we

underlined in the text. It is not clear why the additional sentence (“*Every time someone switches it on, I go to my room and read*”) was added at the very end. We felt it may create unnecessary confusion about the intent of the indirect message since that last sentence is actually literal and representative of the speaker’s behavior. Therefore, in order to follow a more consistent pattern and make the focus of the non-literal utterance (target sentence) be the last sentence, we decided to switch the order of the last two sentences of the dialogue. We kept the same information and clarification (i.e., when someone turn the TV on, I go to my room and read) but reversed the order.

Overall, this translation process proved to be lengthy and at times frustrating since some items were hard to translate into easily accessible dialogues. Since Taguchi did not provide any of the fillers used in her study, we had more flexibility in the creation of these “literal/direct meaning” items.

### 5.1.3. Fillers

Because the test item file we received from Taguchi et al. (2016) did not include the fillers that were used in Taguchi et al.’s study, we had to create our own filler items. The authors indicated that their fillers were dialogues in which the last comment illustrated a direct literal meaning.

With the goal of collecting data that would allow for a possible comparative analysis between direct (or literal) and indirect meaning, we purposely designed our fillers to mirror—or closely relate to—some of the situations that presented non-literal comments. The idea was that the fillers would either end with a sincere comment that was the opposite of an indirect comment used in the non-literal dialogues, or that it would include a similar comment as the non-literal version, but that was to be interpreted as having a direct meaning.

Below is a list of fillers and related non-literal dialogues, followed by an example (Fig. 6) illustrating what we just mentioned (cf. [Appendix C](#) for the full dialogues):

- Filler 1 → IRO 4
- Filler 2 → IRO 6
- Filler 3 → IRO 6
- Filler 4 → IRO 7 (and may also be compared to IRO 2)
- Filler 5 → IRO 13 (and may also be compared to OPI12)

- Filler 6 → IRO11
- Filler 7 → OPI6
- Filler 8 → OPI8
- Filler 9 → OPI2
- Filler 10 → IRO16

It is not that surprising that most fillers were meant to be opposites of ironic comments or indirect opinion (often hidden criticism): these two categories presented the best ground for a reversal of intent that would not be as obvious as a direct vs. indirect refusal. Additionally, Taguchi et al. found that their learners had the most difficulty with irony items which we thought warranted a closer investigation.

Let us look at the French version of our Irony dialogue IRO6. We used this item as the basis to created two fillers (Fill2 and Fill3). First, we will analyze our last statements in bold:

### Figure 6

#### *IRO6 - dialogue*

IRO6dîner
<p>Conversation entre deux amis, Olivier (homme) et Marie (femme), au sujet du repas que Marie a préparé.</p> <p><i>Marie:</i> Olivier, tu veux du dessert? J'ai fait une tarte aux fraises.  <i>Olivier:</i> Ah? Oui bien sûr. J'adore la tarte aux fraises.  <i>Marie:</i> Et qu'est-ce que tu as pensé du poulet-frites que j'ai préparé?  <i>Olivier:</i> <b>Ben en fait, ton repas était vraiment horrible. J'ai presque rien mangé!</b></p>

The intended ironic meaning of Olivier's last turn above (in bold) was meant to be interpreted as a positive statement: he actually ate all his food (also implying that he really likes Marie's cooking). In the video, the male speaker shows his empty plate as a visual cue that he is being ironic when uttering his last sentence.

We created two sets of fillers focusing on a similar situation: the French version of Fillers 2 and 3 are presented side-by-side below.

## Figure 7

### Filler 2 and Filler 3 - dialogues

Filler 2	Filler 3
Conversation entre deux amis, Stéphane (homme) et Natacha (femme), au sujet du diner que Natacha a préparé.  <i>Natacha:</i> Stéphane, est-ce que tu veux un café après le dessert? <i>Stéphane:</i> Ah? Oui bien sûr. <i>Natacha:</i> Est-ce que tu as aimé le diner? Qu'est-ce que tu penses de ma quiche? <i>Stéphane:</i> <b>Ton repas était délicieux et j'ai adoré la quiche. J'ai tout mangé!</b>	Conversation entre une mère (femme) et son fils, Nicolas (homme) au sujet du gâteau préparé par la mère.  <i>Mère:</i> Je suis en train de tester une nouvelle recette de gâteau au chocolat, tu veux goûter? <i>Nicolas:</i> Ah? Oui bien sûr. <i>Mère:</i> Alors qu'est-ce tu en dis, tu penses que ça va plaire à ton père? <i>Nicolas:</i> <b>Oh la la, c'est trop sucré, j'aime pas trop. Je pense que tu devrais mettre moins de sucre!</b>

While Filler 2 is much closer to IRO6 in terms of situation, all three dialogues revolve around food and a potential criticism of someone's cooking skills (positive or negative—in bold). IRO6 presents a seemingly negative literal comment (“**J’ai presque rien mangé**”: I barely ate anything) that has a positive intent (ironic indirect meaning “I ate it all, your food was delicious”). Filler 2 presents a positive literal comment (“**Ton repas était délicieux**” “**J’ai tout mangé**”) with a positive direct intent, that is almost the exact opposite of the ironic comment in IRO6. In the case of Filler 3, we can argue that the negative literal comment (with a negative intent: “**j’aime pas trop**”=“I don’t like it”) is still fairly similar to the comment used in IRO6: in appearance, both seem to offer a negative criticism of the food that they tried, however the speaker in IRO6 is actually being positive (ironic praise) while the speaker in Filler 3 is offering a negative criticism (literal criticism).

#### 5.1.4. Feedback and final modifications

Particular care was taken to maintain length of dialogue consistency across the various types of indirect meaning dialogues as well as the fillers “to control impact on short-term memory” (Taguchi et al., 2016). Below is a table showing the mean and standard deviation for the French dialogue word counts and the video length per categories and overall (the individual item data can be found in [Appendix D](#)). Note that Taguchi et al. only reported an approximate range for the dialogue word count, and did not include video length. Tables [6](#) and [7](#) below provides the

overall average word count for our dialogues and videos as well as the mean for each category of test items.

**Table 6**

*Word count for the French Dialogue*

Item categories	<i>Mean</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>
All included	46.16	5.88	30	56
Indirect refusals	44.25	4.80	35	54
Irony	48.75	5.89	39	56
Indirect opinions	44.56	6.35	30	53
Fillers	47.6	5.50	40	56

**Table 7**

*Length of videos (in seconds)*

Categories	<i>Mean</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>
All included	14.22	2.77	9	24
Indirect refusals	13.00	2.00	10	19
Irony	16.38	2.96	12	24
Indirect opinions	13.44	2.66	9	19
Fillers	14.00	1.83	11	17

Our dialogues were a little longer than Taguchi et al.'s (they reported between 38 and 42 words), but the length of our dialogues and videos were fairly consistent across the different categories. We will discuss the creation of the videos in the next section but it is important to understand that our decision not to alter the dialogues further to reduce the differences in length was mainly due to the fact that, while the irony items and fillers were slightly longer, overall the video clips we ultimately produced were between 9 and 24 seconds ( $M = 14.22$ ,  $SD = 2.77$  across categories), which we thought was a length that would have a limited impact on participants' short-term memory (which Atkinson & Shiffrin (1968, 1971) determined to be around 15-30 seconds).

Before we started filming the videos, we sought feedback from four French native speakers who live in France (two in their mid-sixties and two in their mid-thirties, the latter being the actors in our videos). For their study, Taguchi et al. gave a paper-and-pencil version to four native speakers so they could check for accuracy rates and plausibility of the conversations. Our process was obviously a little bit different, since we used Taguchi et al.'s existing study as our starting point and translated their items. Still, we decided to also confirm the plausibility of the dialogues with our four native speakers. An electronic version of the French dialogues was sent to them along with an informal "opinion survey" (a sample of which can be found in [Appendix E](#)). Reviewers were asked to rate the scenarios on a scale of 1 (not at all plausible/credible) to 5 (completely credible/plausible), which yielded the following results (see [Appendix F](#) for complete data):  $M = 4.76$ ,  $SD = 0.49$ , range = 3-5.

Additionally, we also asked that our reviewers provide feedback on the dialogues to ensure our translation reflected the way French natives really speak. We asked that the four reviewers make notes of all the changes they would recommend and scheduled videoconferences to discuss their comments and suggestions.

Below, we will present some examples of additional modifications that were made based on our reviewers' feedback:

- Minor alterations such as verb conjugations (e.g., changing "*Je préfère...*" to "*Je préférerais...*" in dialogue IRF4 and IRF16, which sounded more natural to our native speakers).
- Adding adverbs or expressions (e.g., "*je n'ai pas très envie d'être exposée au soleil*" in IRF5, or the addition of "*franchement*" in the last turn of dialogue Fill8) for a more naturalistic-sounding language.
- Adding or removing the "*ne/n*" in negative structures. On this particular point, our two older native speakers seemed to disagree with the two younger reviewers: the former wanted to consistently add the first half of the negation, while the younger reviewers felt it was unnecessary given that this was a reflection of oral language. Since the younger reviewers were the ones who would act out the dialogues (and they were also closer in age to our study participants), we followed their recommendation over that of the older reviewers. Low retention of negative "*ne*" in native French is also supported by research from Ashby (1981), Armstrong (2002), Meisner, Robert-Tissot, & Stark (2015), to cite only a few.

- Some of the greetings at the beginning of certain dialogues were also discussed as being unnatural. Based on that feedback we shortened some of them. For example, the first two turns in dialogue OPI9 were translated as follows in our first draft, but the parts that are crossed out below were removed from the final version:

*Julien: Salut, Amélie! ~~Comment~~ ça va? Quand est-ce que tu as ton entretien?*  
*Amélie: Salut Julien! ~~Ça va, et toi?~~ Mon entretien? C'était lundi dernier.*

- A couple of additional alterations were made to individual lexical items. For example, the table became a desk in OPI11 for increased plausibility, “*acheté*” in IRF2—which we had originally used due to its accessibility for lower-level learners—was replaced by “*commandé*”—which is more commonly used in French in that context. Sometimes, whole sentences were edited so that the dialogues would sound more credible to French native speakers. For example, in dialogue IRF11, a couple is having dinner, and at the end of the conversation, the husband offers his wife strawberries or chocolate cake for dessert (“*J’ai des fraises ou du gâteau au chocolat.*”). The wife’s answer in our first draft based on Taguchi et al.’s version (“*Je n’aime pas les fraises*”) appeared too abrupt to our reviewers and very unnatural. The explanation provided by our reviewers was that the husband would have likely known that his wife did not like strawberries, but her statement made it sound like she was providing him with a fact he was not aware of. This last turn was then edited to: “*Ben tu sais bien que j’aime pas les fraises,*” which was meant as a reminder to the husband. In both cases, the intended indirect refusal and implication (I do not want strawberries for dessert: I want chocolate cake) remained unchanged.
- Similarly, some elements were added to some dialogues to help with comprehension. Dialogue IRO15 features a conversation between two roommates. Antoine just bought new chairs at the store and Clara is not too happy with them. When she tries one, it wobbles and she sarcastically comments that “they don’t look old at all” (meaning that they do look old). Due to the apparent confusion that the scene created in our reviewers we discussed the fact that Antoine had likely bought second-hand furniture. To remove any ambiguity, the expression “*magasin d’occasion*” (second-hand/thrift store) was used in our final version.
- In dialogue Fill7, the conversation happens between two friends who have not seen each other in a long time. In our original dialogue, the female speaker commented on the male speaker’s hair color. Our reviewers thought it would make more sense to reverse

the roles and have the male speaker inquire about the female's speaker changing her hair color from brown to blond. They argued that men in their thirties in France rarely dye their hair. After we made that switch, our female reviewer (who would be acting the scene) realized that since her hair was brown at that time, it would therefore be more accurate (and facilitate comprehension) if the dialogue referenced her hair changing from blond to brown. As a result, we also incorporated that change in our final version.

There were also instances when the changes suggested by our native speaker reviewers could not be implemented because it may have interfered with the learners' ability to understand (for example, two native speakers recommended switching from "libre" to "disponible" in IRF6 but we were not sure whether that was a word that our intermediate-level learners would comprehend) or because it went against the design of our study (and the original study by Taguchi et al.). The most significant example of that was found in the indirect refusals: at least three of our native speaker reviewers consistently indicated that it seemed awkward that the person would not respond directly to the questions in these dialogues. They recommended adding an IFID (cf. [section 2.7](#)) in the form of a negative answer in the last turn, which would have rendered that dialogue irrelevant for the purpose of the current project, as the refusal would then change from indirect to direct. This brings up a very important and valid point regarding cultural norms: while we followed Taguchi et al.'s design and purposely did not use explicit negative words (*Non, non merci*, etc.) in the indirect refusals, we previously mentioned ([subsection 4.1.1](#)) that the conventional pattern identified by Taguchi and other authors (providing a reason or excuse for the refusal) may not sound as natural in French (as it is in Spanish). It is possible that French people's preferred way to refuse an invitation is to explicitly (yet succinctly) reject it (using negative words like "Non" or "non merci") before providing their reason or justification or possibly offering an alternative (as seen in Taguchi et al., 2013).

Finally, some of the issues that we had encountered when we first translated the dialogues were also discussed with the four native speakers, in an attempt to get additional suggestions for improvement. However, at times, the consensus was to leave the dialogues as they were, since no perfect solution was found. We suspect that this issue and the awkwardness of some of the situations may have also been present in the Spanish dialogues of the initial study. As a matter of fact, upon asking a few Spanish speakers, they confirmed that they were not sure whether they understood the last part of certain dialogues. Here are a few examples of these problematic dialogues:



- Dialogue IRO12 focuses on the interaction between a husband and his wife. Upon being asked whether she would like some coffee in the morning, the wife replies “*Regarde-moi: je suis complètement réveillée là!*” (which was a more or less direct translation of the Spanish version “*¿Qué te parece, cariño? ¡Estoy tan despierta!*” which roughly translates from both versions to “Look at me (honey): I am so wide awake!”). As an ironic comment, the statement was meant to be interpreted as a positive answer: the wife definitely needs coffee because she is not completely awake yet. However, our reviewers did not understand this statement and were confused by it. We decided to keep it as such because it replicated the original version fairly closely. Since Taguchi et al. did eliminate some of the items based on how their group of native speakers answered them (taking into account accuracy and/or speed), we hypothesized that this dialogue may be removed from our final version if native speakers who participated in our pilot study were to have trouble with it.
- Dialogue IRF3 presented similar challenges but once we discussed it with our native reviewers, they understood the indirect meaning better. The main problem with that dialogue was the reference to a “*veste*” or jacket in the last sentence “*Je n’ai même pas l’argent pour m’acheter une veste.*” (“I don’t even have money for a jacket”). Our reviewers thought that the item of clothing seemed like a very random object to refer to. The intended meaning of this last sentence was an indirect refusal following an invitation to travel to Europe on the basis that the female character did not have enough money.
- IRO16 presented some challenges for translation and comprehension. In Taguchi et al.’s version, we are introduced to two friends who are at a coffeeshop ordering drinks. The female character orders a “double chocolate cappuccino.” The following ironic statement (from the male speaker) is intended as a comment regarding the fact that the female speaker ordered a very sweet drink (“*¡Qué amargo que tomas tú el café!*”/“You like your coffee a bit bitter, right?”). The concept of drinking “bitter coffee” (or “*café amer*”) was problematic to our French native reviewers who explained that bitterness did not necessarily exclude sweetness in their opinion. We discussed the possibility of using the French “*café noir*” (“black coffee”) instead, but the consensus was the same: in France, one can order a black cup of coffee but still add sugar to it. Their view was that “black coffee” just signified that no milk or cream was added but it did not omit the possibility of adding sugar. The only solution and compromise we could therefore come up with was to use a negative statement combined with an opposite lexical item (i.e.,

from “bitter coffee” to “sugary drink”): “*Ah ben je vois que tu n’aimes pas les boissons sucrées toi, non?!*” (“Well, it sounds like you really don’t like sugary drinks, now do you?”)

- Another dialogue that we thought may cause some issues was IRF13. The last sentence is a comment from the male speaker regarding the female speaker’s spending habits (and the fact that she wanted to buy another pair of shoes while she had just purchased one): “You shouldn’t spend more money on shoes.” (In Spanish “*No deberías gastar más en zapatos.* »). We translated as « *Tu ne devrais pas dépenser encore plus d’argent pour des chaussures.* ». Our first impression was that the sentence was awkward due to the combination of a negation and double adverb “*encore*” and “*plus.*” We did not anticipate that our native reviewers would find its comprehension challenging. Without being able to provide an explanation as to why they found this wording confusing, they initially proposed “*tu ne devrais pas dépenser trop d’argent.*” However, we agreed that it did not convey exactly the same idea. After discussing alternate possibilities (including removing “*encore*”), we decided to keep the sentence the way it was.

#### 5.1.5. Videos

Two French native speakers were recruited for the filming of the videos: one female and one male in their mid-thirties. Both live in the South-West of France but do not have a very strong accent. Taguchi had indicated that the scenes were filmed in one of the researchers’ home (in the same room) and that they had made limited use of props depending on the dialogues (e.g., “some pieces of clothing for a situation that involved dresses, and jackets and books for others that dealt with going to school and talking about the cold weather outside.”-Taguchi, personal correspondence, April 2019). She also explained that while their actors were not coached, “two of the authors in [their] paper agreed on certain intonation cues, as well as the use of gestures and facial expressions to help participants understand the situations and speech” (Taguchi, personal correspondence, April 2019).

Regarding the location, we agreed to film all the videos in the same house. However, different scenes took place in different rooms for more credibility. While we wanted to avoid the possibility that location became a confounding factor, we also thought that having all the videos filmed in the same room (a living-room like in Taguchi et al.’s study) would actually be more problematic when it came to comprehension. We wanted to make sure the setting was somewhat

realistic based on the content of the dialogues. Another aspect that also contributed to that decision was the possibility that participants would experience survey fatigue, and we thought that using different locales may reduce it a little bit. Some scenes were filmed outside, some were done in the living-room/kitchen area (all the dialogues revolving around food, or the ones took place in a restaurant/cafeteria etc.). Conversations that were set in a classroom, between students working on assignments or at the university, were filmed at a table or desk to increase credibility. Props that were used included items of clothing and accessories, books, plates, mugs and eating utensils, all in relation to the dialogues for which they were used. In some instances, the props were deemed absolutely essential to the comprehension of the indirect meaning (e.g., IRO6 when the male speaker says he barely ate anything but shows an empty plate).

Our two speakers were not professional actors and were briefed on the general purpose of the study. They were instructed to play the scene in a way that was as natural and authentic as possible, but to pay attention to gesture, intonation and facial expressions. However, we did not give them any instructions as to which specific gesture, intonation or expressions to use. They knew that the videos would be viewed by non-native speakers, and consequently they had to use natural but clear enunciation and speak at a normal pace. The filming took place on four different days. After we reviewed the videos filmed on the first day, we made additional recommendations: videos should be filmed in landscape mode rather than portrait. Indeed, since our recommendation was for participants to watch the video on a computer screen, the horizontal format worked better than the vertical one. In addition, we indicated to our filming crew that it would be best to film the remaining videos indoors so as to avoid issues with outdoor noise interfering with the dialogues (for example, two videos had to be redone due to the wind being too loud, or a plane/tractor passing by). Once all the videos were filmed, a second viewing revealed some additional minor issues that required a second take of a couple of clips (e.g., the person who filmed the video could be seen in the reflection of a mirror or in the oven in two videos, or one video did not show the face of the male speaker in a way that would help with comprehension). Overall though, most videos were judged to be of excellent quality and the acting deemed natural yet accessible for non-native speakers.

As shown in the previous section, the average video length was between 13 and 16.38 seconds across categories. Given that our final study was to include 48 videos (including two practice items), that would equate to a little over 11 minutes' worth of video viewing, which seemed like an appropriate amount for an activity that was scheduled during a one-hour class period.

All videos were made using an 8-megapixel handheld video-capture device, capable of recording images at up to 3264 by 2448 pixels. Using high definition videos provided greater quality that would guarantee that our participants would not miss important details during the video-viewing phase of the experiment. Most of the video files we saved were around 19 MB (with the largest file at 24.9 MB). However, when we tried to upload the videos to our online survey platform, Qualtrics, (see additional details in [sub-section 5.2.2](#)), we found out that files larger than 16MB could not be uploaded to the platform library. Therefore, the majority of our videos had to be compressed to a smaller size (around 15 MB) before being stored in the Qualtrics library. A free online video converter website (<https://www.onlineconverter.com/compress-video>) was used to that effect.

Two videos (IRO10 and OPI13) had to be created from merging two files: these two conversations were filmed in two separate takes based on the situational indication to accommodate for a time lapse or a change of location (i.e., the second half of IRO10 happens “*One hour later - When Emmanuel comes back home,*” and in OPI13, the characters are first in their home and then “*at the store*” in the second half). The video editing software ShortCut was used to combine the files and the situational prompt (in italics above) was added as a transition between the two clips of each video (in both English and French).

Once all the videos were created, we turned our attention to the development of the online survey. Following the example of Taguchi et al, we wanted that instrument to provide an easy way for participants to watch the video, answer the multiple-choice questions, as well as offer all the options we needed to collect the data we wanted to gather. In the following we will detail the process by which we created the multiple-choice questionnaire as well as the online survey<sup>42</sup>.

## 5.2. Instrument development

The following sub-sections will focus on the creation of the multiple-choice questionnaire for the replication, as well as the design of the computerized online survey. We will make sure to

---

<sup>42</sup> NB: The tasks and videos for our study (video files, dialogues and multiple-choice questionnaire) have been uploaded to the IRIS database and are freely accessible: <https://www.iris-database.org/details/j7a8J-AnCBC>

discuss the challenges we encountered and the solutions we adopted for the final instrument.

### *5.2.1. Multiple-choice questionnaire tasks*

While some of the multiple-choice answers created by Taguchi et al. were problematic (see [Chapter 4](#)) a true replication study required us to keep the system that was used in their experiment. Since we used Taguchi et al.'s dialogues as the basis for creating our videos, it made sense to also follow the same pattern for the multiple-choice options. As a reminder, Taguchi et al. (2016) applied the following principles in order to create their answers (p. 687):

- #1: The first option was the CORRECT answer. For the experimental items, that meant the intended meaning, which was different than the literal meaning expressed by the last statement in the conversation. For the fillers, the correct answer was similar to the last statement (paraphrasing the literal meaning).
- #2: **Principle 1** was designed to be the most remote from the intended meaning (or the “opposite to the target indirect meaning”). In the case of an ironic comment using a meaning reversal, we would expect these to be very similar to the last turn of the conversation.
- #3: **Principle 2** option would “contain words taken from the last part of the conversation.”
- #4: **Principle 3** would offer a statement that related to the overall conversation.

The MCQ questions were presented in English to the learner group and in the target language to the native speaker group. Therefore, we also needed to translate the MCQ into French for our native speaker group (where Taguchi et al. used Spanish). In our first draft, we translated Taguchi et al.'s statement as closely as we could. After receiving feedback on the dialogues and making some modifications, we had to also alter the multiple-choice answers to reflect the changes made in the dialogues, as exemplified in [Figure 8](#).

**Figure 8**

*IFR1 – task comparison between Taguchi et al. (2016) and our version*

Taguchi et al. (2016) – <b>IRF1cena</b>	Our version – <b>IRF1repas</b>
<p>This is a conversation between Carlos (male speaker) and Adriana (female speaker). They are making plans for tonight.</p> <p><b>SPANISH:</b>  <i>Carlos:</i> ¡Hola, Adriana! ¿Tienes planes para esta noche? ¿Te gustaría ir al cine hoy?  <i>Adriana:</i> Pues no creo. Tengo mucho por hacer.  <i>Carlos:</i> ¿Y qué tal ir a cenar? No nos llevará mucho tiempo.  <i>Adriana:</i> <u>Tú siempre tardas mucho en comer.</u></p> <p><b>English translation:</b>  <i>Carlos:</i> Hi, Adriana! Do you have plans for tonight? Wanna go to the movies?  <i>Adriana:</i> I don't think so. I have lots of things to do.  <i>Carlos:</i> And what about getting dinner? It won't take us that long.  <i>Adriana:</i> <u>It always takes you forever to eat.</u>            (42)</p> <ol style="list-style-type: none"> <li>1. Adriana doesn't want to go out for dinner tonight with Carlos. (CORRECT)</li> <li>2. Adriana is going to have dinner with Carlos tonight. (Principle 1)</li> <li>3. Adriana thinks waiters at the restaurant are <b>slow</b>. (Principle 2)</li> <li>4. Adriana thinks she has time to have dinner with Carlos. (Principle 3) (38)</li> </ol>	<p>This is a conversation between François (male speaker) and Cécile (female speaker). They are making plans for tonight.</p> <p><b>FRENCH</b>  <i>François:</i> Salut Cécile! Qu'est-ce que tu vas faire ce soir? Tu veux aller au cinéma?  <i>Cécile:</i> Non, je peux pas. J'ai beaucoup de choses à faire.  <i>François:</i> Tu veux diner avec moi? Ça ne prendra pas trop de temps.  <i>Cécile:</i> <b>Tu manges pas assez vite.</b></p> <p><b>English translation :</b>  <i>François:</i> Hi, Cécile! What are you doing tonight? Wanna go to the movies?  <i>Cécile:</i> I can't. I have lots of things to do.  <i>François:</i> Do you want to have dinner with me? It won't take us that long.  <i>Cécile:</i> <b>You don't eat fast enough.</b>            (39)</p> <ol style="list-style-type: none"> <li>1. Cécile doesn't want to have dinner with François tonight. (CORRECT)</li> <li>2. Cécile is going to have dinner with François tonight. (Principle 1)</li> <li>3. Cécile thinks waiters at the restaurant are <b>fast</b>. (Principle 2)</li> <li>4. Cécile thinks she has time to have dinner with François tonight. (Principle 3) (37)</li> </ol>

In the dialogues above, notice the modulation in the last statement of the conversation (the move from the Spanish “*Tú siempre tardas mucho en comer.*” to the French “*Tu manges pas assez vite*” was a recommendation from our four reviewers who agreed that it sounded better to their native speaker ears), which also impacted our third option (as we were applying Taguchi et al.’s Principle 2 according to which option #3 should contain a word taken from the last part of the conversation: “fast” which refers to “vite” in the dialogue). It is important to note here that the issues regarding these principles that we mentioned in [section 4.2](#) were present in this

particular dialogue and by switching from “slow” to “fast” we essentially fixed the problem by applying the principle designed by Taguchi et al.

To be able to run quantitative analyses on the response speed (which in both Taguchi et al.’s and our study was measured from the time the answers appeared on the screen to the time participants selected their answer), we needed to ensure that the multiple-choice options were all around the same length. The assumption was that it would take a similar amount of time for participants to read text that contains a similar number of words. We realize that this is only an approximate method of assessing speed but we followed the system used by Taguchi et al. They reported a range of 34 to 41 words across all items in English (Taguchi et al. 2016, p. 687 “mean = 38.30, *SD* = 2.60”). The average length of our multiple-choice options in English (ranging from 32 to 43 words) was slightly below that of Taguchi et al., but with a slightly smaller standard deviation as shown in the table below. Some adjustments were made to the fillers that were originally significantly shorter than the rest of the items. Taguchi et al. did not report response time for fillers in their paper but in our own analysis, we wanted to be able to compare response time for fillers with response time for experimental items. We therefore, made sure the mean word counts for all item categories were similar. Particular attention was given to the pairs of fillers and corresponding experimental items so they could be of equal length. The same principle was applied to the creation of multiple-choice options in French (for the native French speakers), so we could compare their response time to that of our L2 learners. [Table 8](#) presents the descriptive statistics concerning word length for our MCQ items in both English and French. Means lengths of MCQ items in French and English are very similar, which will allow for an easier comparison.

**Table 8***MCQ word count (English and French)*

Categories	English MCQ word count			French MCQ word count		
	Mean	<i>SD</i>	Range	Mean	<i>SD</i>	Range
All included	37.78	2.53	32-43	37.88	2.48	31-43
IRF	38.13	2.09	32-43	38.13	2.09	35-41
IRO	38.00	2.07	34-41	38.19	2.20	34-41
OPI	37.44	3.39	32-43	37.38	3.26	31-43
Fillers	37.40	2.50	32-43	37.80	2.25	33-40

### 5.2.2. Online survey questionnaire

Our online survey was created using Qualtrics<sup>XM</sup>. While this product is commonly used throughout the US, the institutional subscription-based version we had access to presented some limitations. We did not have access to any in-person support and the bulk of our design process was done by trial-and-error and searching for answers on the Qualtrics Community forums. Qualtrics rely on:

- a user-friendly interface with rich content editor, different question options (multiple choice, text entry, matrix tables, etc.), and the ability to customize the look and feel of the survey, and different ways to distribute it.
- an html and a Java Script box for users to code additional functions and features.

Two final versions of the surveys were created: one designed for our L2 learners and one for the native French speakers. Additional information (e.g., consent) was also required in the English version (see [sub-section 7.1.3](#)).

#### 5.2.2.1. Presentation of the items

The English version of the survey was first created. Video files were uploaded to the Qualtrics library. After the consent page (which we will discuss in the next section), a demographic page was added to collect biographical data and information needed to assess participants' language level and exposure (particularly for the L2 learners) (see [Appendix G](#)). Next, a page of



directions (see [Appendix H](#)) presented the following instructions (and information—summarized below):

- 1) Watch each video carefully.
- 2) Click on next to access the multiple-choice answer.
- 3) Select the appropriate answer.
- 4) Participants were told that the last step would be timed and that they should try to stay focused (but take the time they need to answer properly).

Following the instructions, two practice items were presented in the same fashion as the rest of the test. The only difference, for our purpose, was that these practice items were not randomized. They were selected amongst the experimental items that were discarded after the pilot test was administered to our native speaker group.

A brief description of the scene appeared at the top of each video which was sized for optimal viewing on a wide-screen standard-sized computer screen (width = 1000 for horizontal format video and height = 800 for vertical format). After watching the video, participants clicked on the “next” button to access the next page which corresponded to the MCQ presenting four options in a randomized order. Each item (video and associated MCQ) was also presented in a randomized order. We ran into a few challenges when trying to use the timing feature available on Qualtrics along with the randomization option. Solutions will be presented in our next section.

The French version was created by cloning the English version: all the instructions, video descriptions, and MCQs were translated to French.

#### 5.2.2.2. Challenges and solutions

Taguchi et al.’s study was designed so that the survey would automatically advance from page to page. Qualtrics did not offer as much flexibility and we wanted to have the option to stop or pause the survey in case technical issues arose when participants were taking the survey. A timer could have been placed on pages containing videos so that the next page would appear as soon as the video was over (and each individual page would have had to be set to the exact time of each video), but it occurred to us that this solution could be extremely problematic should the video take longer to load, or should participants run into other computer issues that required

them to pause the video or to stay on the page longer than anticipated. Taguchi et al. had indicated that participants were automatically taken to the MCQ, which meant they only had the opportunity to watch each video once. That option was not available to us in Qualtrics. Therefore, we stipulated that participants should only watch the videos once in our instructions. At the end of the survey we also added a question requesting participants to confirm whether they had only watched each video once or not. We were fairly confident that it would not be an issue given that participants had a limited amount of time to complete the task (one class period: approximately one hour), and that the stakes were not so high that they would want to cheat and not follow the rules (American participants received a completion grade for the activity). Additionally, the fact that we were going to be present in the computer lab during the survey administration to our L2 learner group could also be a dissuasive factor. Similar instructions were provided to the native speaker group; it was unlikely they would need to watch the videos more than once to understand the dialogues.

Participants in Taguchi et al.'s experiment used the number pad to select their answers. Qualtrics did not offer an easy way to implement this answer selection method: instead we used a mouse click option. However, just like with Taguchi et al.'s survey, once the participants selected one answer, the MCQ would disappear and the next page would come up. We used the survey auto-advance feature (which was only applied between the MCQ page and the next video). The "back" buttons were also disabled so as to prevent participants from rewatching the video or editing their answers later.

Because videos and MCQ needed to be presented on separate pages, but each video was to be associated with the corresponding MCQ, the randomization of items required that individual blocks be created for each pair of videos and MCQ. We designed a block template that we then used for the creation of all items. Between each page showing a video and the following page containing the MCQ answers, a "timing" function was inserted to keep track of the time lapse which was recorded between the time the MCQ page loaded and the mouse click selecting the participant's answer. Another reason why blocks needed to be used was that the timing feature could not be used in conjunction with the randomization of questions. Therefore, the randomization was selected at the Survey Flow level.

We asked several volunteers (including the four native speakers who had served as reviewers in the initial stages of our test development) to test the platform and notify us of any issues. A

couple of adjustments were made based on their feedback (e.g., dialogues that did not match the video and had to be switched). We also realized that while we did not have any issues loading the videos on our institution's campus, some of our testers were not able to access the video clips in a timely manner. Therefore, we used the same online website to compress the files and reloaded them to both test libraries (L2 and native speaker).

Once the instrument was created and finalized, we moved to the first phase of our data-collection plan: the pilot study. [Chapter 6](#) presents the methods and results of our native speaker study as well as an explanation of the process we used to identify items that were to be removed for the final version of our test (the L2 learner test).

## Chapter 6 - Experimental study - Pilot with native speakers

Taguchi et al. conducted their pilot study with native speakers for the purpose of testing and improving their instrument. The results were used as part of a task selection process: those items that were deemed more difficult (either due to a low accuracy score or a high response time or both) were eliminated from the L2 study. Because the pilot survey was quite long (16 items x 3 indirect categories + 10 literal fillers = 58 items), Taguchi et al. removed four items per indirect meaning categories (a total of 12 experimental items were discarded) to reduce the total number of items to 46 (12 items x 3 categories + 10 fillers = 46 items). We followed the same procedure.

The goal of our pilot study was two-fold:

- 1) To evaluate native speakers' performance on the test, specifically with the goal of identifying potential differences in comprehension between the different categories of meanings, and in order to compare them with our learner group and Taguchi et al.'s native speaker group;
- 2) To select the 12 experimental items to be removed.

Therefore, in the section below, we will present our methods for the administration of our pilot study ([section 6.1](#)). In a second part ([6.2](#)), we will report our findings on the impact of the different meaning categories based on the statistical analysis of the native speaker dataset. Finally, we will present a summary of the results and describe our process for identifying the items that will be removed from the pilot study, following Taguchi et al.'s protocol ([section 6.3](#)).

### 6.1. Methods

This section will concentrate on the native speaker participants and procedures for administering the pilot test to this group.

#### 6.1.1. Participants

Participants for the pilot study were recruited via an email announcement (see [Appendix I](#)) sent

to students at two medium-size public universities in France. The requirement to participate in the study was that students should be native French speakers. No compensation was offered to participants, who answered the questionnaire on a volunteer-basis. A total of 41 participants accessed and started answering the online survey, but only 29 completed the entire questionnaire. Incomplete answers were discarded and only complete sets were kept for the purpose of this study. The average age of our native speaker participants was 23.07 years ( $SD = 4.55$ ) ranging from 19 to 43 years old. The majority of the participants had already completed or were in the process of finishing their undergraduate degree, or “*licence*” in French (average of 3.48 years of higher education;  $SD = 2.42$ , range = 0-5). One participant reported their education level as “*baccalauréat*” (equivalent to the high school diploma). However, it is possible they only considered their highest degree in hand but were actually pursuing a bachelor’s degree.

### 6.1.2. Procedures

The recommendation made to the volunteers who completed the online survey was to use a computer to access the survey. It is however unclear whether participants used a computer or a mobile device. Regardless, we did not anticipate that using a mobile device would have a negative impact on native speakers’ comprehension. The Qualtrics survey was also designed for enhanced mobile compatibility, which we tested in the event that native speakers chose to opt for that format. We did not notice any issues with the mobile delivery.

Participants all started and finished the survey the same day. Our data collection window was from October 10 to October 23, 2019. Since timing was only recorded when participants were selecting the multiple-choice answers, it was possible for them to stop the survey when they landed on a video page. The total amount of time that participants spent on the survey varied (range = 27 minutes to 139 minutes, mean = 46.25,  $SD = 27$ ). After removing the three outliers that accounted for the highest amount of time (139 min., 74 min., and 56 min.), the average time spent on the survey was about 37 minutes ( $SD = 7$ ) which is consistent with what we had anticipated (our recruitment message had announced a 35-minute survey). We hypothesize that the participants who took longer to finish the survey may have needed to stop to tend to other things before returning to the survey. In no way did this issue seem to have negatively impacted the response time results however (results will be discussed in our next section, but average response time for each multiple-choice question across the survey was 7.37 seconds, with a

standard deviation of only 1.73 seconds).

In the next section we will present the descriptive analyses of native speaker accuracy scores and response times and discuss potential differences with native speakers' performance in Taguchi et al.'s study.

## 6.2. Native speaker performance and impact of indirect meaning on comprehension

In their paper, Taguchi et al. reported on their native speaker group performance. They however explained that the focus of their study was learner comprehension, and we are meant to infer that the native speaker results were only provided as a point of reference. Nevertheless, we also wanted to include the results of our native speaker performance because we felt that: 1) it was important for us to see if the different categories of indirect meaning had the same impact on the native speaker group in Taguchi et al. that they did on our native speaker group; 2) we wanted to be able to compare our native speaker group to our learner group the same way Taguchi et al. did.

In the next three sub-sections, we will present the analysis and results for the comprehension of our test items by the native speaker group. We will then discuss the differences between Taguchi et al.'s native speaker findings and results from our native speaker group. A discussion on the differences between our native speaker group and our learner group will be presented in [Chapter 7 \(section 7.4\)](#).

### 6.2.1. Data Analysis

Our pilot study results were exported from Qualtrics into an Excel spreadsheet. Incomplete sets of answers (12 participants who did not finish the test and only provided some of the answers) were removed from the spreadsheet and we only kept the 29 participants with a complete set of answers.

In order to present a descriptive analysis of this native speaker set, we looked at average accuracy scores and average response times for each of the native speaker participants. We first converted the text answers (Q<sub>n</sub> answers) into scores, giving a score of 1 to expected answers

and a score of zero to incorrect answers, following Taguchi et al.'s system. Taguchi et al. used a 0 to 12 interval scale for accuracy but since our literal meaning category only included 10 items, we decided to convert the accuracy scores into percentages, in order to include the literal meaning category in our statistical analysis later on. As a result, accuracy was operationalized as a percentage score for each type of indirect/literal meaning.

Comprehension speed was operationalized as response times based on the average number of seconds it took for participants to correctly answer questions in each category. We confirmed that participants only watched each video once (all answered "Yes" to the corresponding question on the questionnaire). Taguchi et al.'s stated that they only used the response times "taken to answer each item correctly" (p. 688), so we removed the incorrect answer data from our response time dataset. Indeed, to better measure the impact of the various conditions on comprehension, we needed to exclude those items that were answered incorrectly, as it meant that participants were not able to access the correct interpretation. Moreover, we decided to use the literal meaning items not only as distractors, but also as part of our analyses, thus expanding upon Taguchi et al.'s study, since they excluded the distractors from their analyses. Since it is generally necessary to understand the literal meaning in order to access the indirect meaning, including the fillers in our analyses would provide a good benchmark for the comprehension of literal meaning so that we could compare those items to the indirect meaning categories.

Using the `pastecs` package in R Studio (2021), we tested both our native speaker datasets (i.e., accuracy and response time) for normality. The results indicated that neither set was normally distributed: using the `stat.desc` command, we found that all four categories in our accuracy dataset presented a negative skew, and the category of irony in our response time dataset was positively skewed and leptokurtic. Therefore, we ran Robust one-way repeated ANOVAs (Field, Miles, & Field, 2012) using the `WS2` package (Wilcox, 2012) in order to determine whether item type (IRO, IRF, OPI, Fill) significantly influences accuracy and/or response times. The `rmanova()` command in this package computes a one-way repeated measures ANOVA for the trimmed means (Field & Wilcox, 2017). Homoscedasticity assumption was not required. Corresponding post hoc tests were performed using the `rmmcp()` command (Field et al., 2012, Mair & Wilcox, 2020).<sup>43</sup>

---

<sup>43</sup> The `rmmcp` function we used for this post hoc test relies on "Hochberg's approach to control for the family-wise error (FWE)." (Mair & Wilcox, 2020) and is based on a 20% trimmed mean (Field et al., 2012)

Effect sizes (both  $d$  and  $r$ ) were calculated using the `compute.es` package in R Studio (2021). In this study, we used Plonsky and Oswald’s (2014) recommendations for the interpretation of effect sizes. According to Plonsky and Oswald, for within-subjects comparison, for  $d$ , a small effect size lies between 0.6 and 0.99, a medium effect size lies between 1.0 and 1.39, and a large size is anything beyond 1.40. For  $r$ , an effect size of 0.25 is considered small, of 0.4 medium, and of 0.6 large.

6.2.2. Results

The result section will be divided into two sub-sections: first we will present native speakers’ comprehension accuracy results and in a second sub-section, we will focus on their comprehension speed.

6.2.2.1. Comprehension accuracy

The results from our descriptive analysis of native speakers’ accuracy is presented in [Table 9](#).

**Table 9**  
*Native Speakers’ Accuracy Scores*

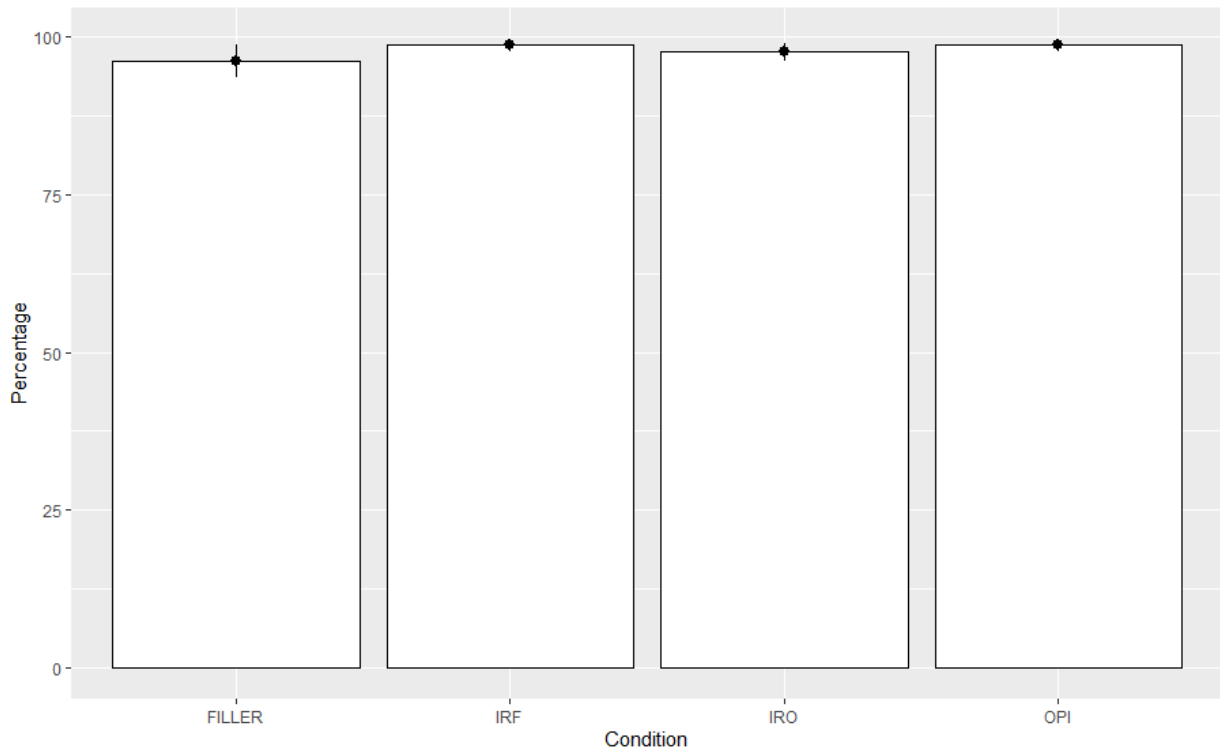
	Mean	SD	Min.	Max.
Indirect refusals	98.71	2.58	93.75	100
Irony	97.63	3.89	87.5	100
Indirect opinions	98.71	2.58	93.75	100
Fillers/Literal	96.21	6.77	80	100
Overall	97.81	3.95	88.75	100

As shown in the above table, native speakers’ overall accuracy was very high (97.81%). Using the packages `ggplot2` and `Hmisc` we created a bar graph with error bars for our accuracy scores based on participants’ averages in each category (see [Figure 9](#)).



**Figure 9**

*NSs' Accuracy Scores – bar graph with error bars*



Based on the above graph, we knew the native speakers' accuracy scores were high and fairly similar across conditions. This was confirmed by the statistical analysis presented below.

The robust one-way repeated measures ANOVA found that the variable of item type did not have a significant impact on the accuracy of responses,  $F(2.12, 38.25) = 0.4865$ ,  $p = 0.6298$ . This means that there were no significant differences between the different types of indirect/direct meaning in terms of comprehension when looking at accuracy for native speakers.

#### 6.2.2.2. Comprehension speed

[Table 10](#) displays descriptive statistics for the native speakers' response times (in seconds) for items to which correct answers were given. These results are based on averages per participant.

**Table 10**

*NSs' Response Times on correct answers*

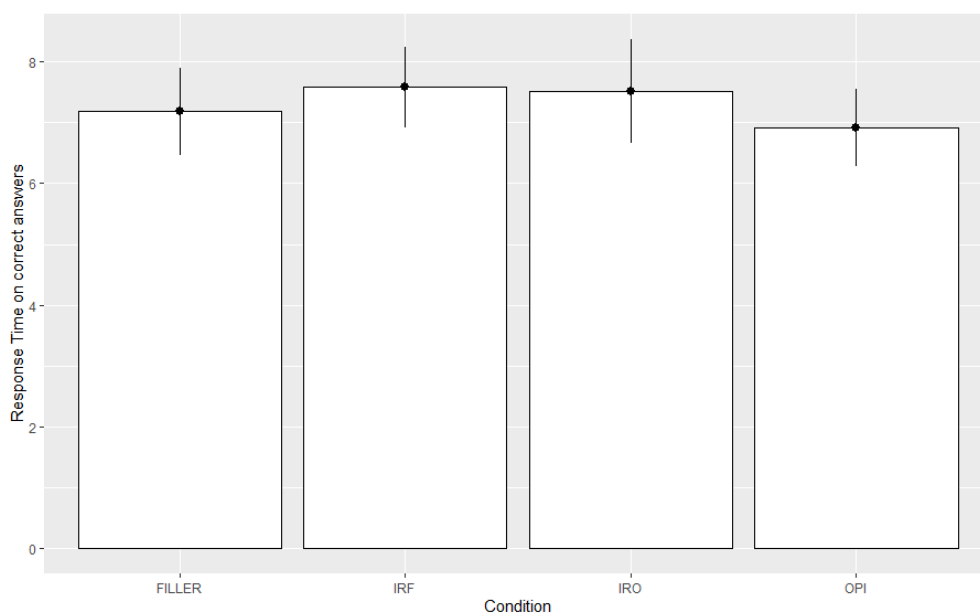
	Mean	SD	Min.	Max.
Indirect refusals	7.58	1.74	4.60	13.58
Irony	7.52	2.25	3.91	13.49
Indirect opinions	6.89	1.67	4.17	12.13
Fillers/Literal	7.18	1.88	3.86	12.87
Overall	7.29	1.88	4.14	13.02

[Table 10](#) shows that indirect opinions resulted in the fastest response time while indirect refusals was the category that took participants the longest time to answer correctly.

Using the packages `ggplot2` and `Hmisc` we created a bar chart with error bars for participants' comprehension speed in each category (see [Figure 10](#)).

**Figure 10**

*NSs' Response Times on correct answers - bar graph with error bars*



The above graph shows a longer response time for indirect refusals and irony (though irony had

a wider error bar) as compared to the indirect opinions and fillers.

Our Robust one-way repeated measures ANOVA revealed a significant difference in native speakers' comprehension speed across all four categories,  $F(2.77, 49.78) = 3.9361, p = 0.0156$ . Our post hoc test showed that response times were only significantly different for the OPI vs. IRF comparison ( $p = 0.00239$ ). Comparisons of IRF vs. IRO ( $p = 0.73059$ ), IRF vs. FILLER ( $p = 0.04335$ ),<sup>44</sup> IRO vs. OPI ( $p = 0.12042$ ), IRO vs. FILLER ( $p = 0.27361$ ) and OPI vs. FILLER ( $p = 0.37617$ ) were not statistically different. The effect size for the comparison between OPI and IRF was very small ( $d = 0.4, r = 0.2$ ).

### 6.2.3. Discussion

Results based on the pilot study show that native speakers did not really experience difficulties understanding indirect meaning since accuracy scores were homogenous across all four categories of meaning (indirect refusals, irony, opinions and fillers/literal items—for a total average accuracy of 97.82%). There was no statistical difference in comprehension accuracy between the different categories. The only significant difference found in comprehension speed was between indirect refusals (slowest time) and indirect opinions (fastest time). A central claim in Taguchi et al.'s (2016) study is that conventional implicatures (in particular indirect refusals) are easier to comprehend due to the predictable pattern of interaction and therefore should result in high accuracy and fast comprehension speed. Our native speaker data does not necessarily uphold this claim (or previous findings regarding conventionality effect –cf. [section 3.2](#)). Conventionality of indirectness was not a factor that facilitated comprehension for our native speakers (lack of conventionality did not hinder comprehension either), which can be expected given their native proficiency. Taguchi et al. (2016) also note that higher proficiency levels (in L2 learners, but that would obviously be true for native speakers) often result in no significant difference between conventional and non-conventional items (p. 694).

One possible explanation for the longer response times observed in the indirect refusal category could be the absence of IFIDs in the dialogue, which may have in fact given pause – perhaps literally – to native speakers. Indeed, in [sub-section 5.1.4](#) we mentioned that some of our native

---

<sup>44</sup> This p-value ( $p = 0.04335$ ) is not significant here due to the multiple comparisons. R Studio (2020) automatically indicates the significance (TRUE or FALSE) based on the following according to Field et al. (2012, p. 578): “If the value of p. value is less than the critical value (p. crit) and the confidence interval does not cross zero then the comparison is significant.”

reviewers from whom we had asked for feedback on the dialogues, had expressed some hesitation with regard to the indirect refusals, and felt that the addition of an explicit IFID would sound more natural. The absence of the IFID did not prevent our pilot study native speakers from reaching an accurate interpretation (given that the accuracy rate was the highest for indirect refusals<sup>45</sup>), but it may have led to longer processing times.

Since our dissertation focuses on L2 comprehension of indirect meaning, we will not discuss these native speaker results<sup>46</sup> in depth. However, we nonetheless wanted to compare them with Taguchi et al. (2016) as part of our replication. While Taguchi et al. explicitly stated that the purpose of their study was not to compare the native speaker performance with that of the L2 learner group, they did include the statistical results of their native speaker datasets. They reported that their native speaker group presented similar differences to their learner group for both accuracy and speed. In terms of accuracy scores, they indicate that (p. 689):

[t]here was a significant difference across three item types, chi-square = 10.67,  $p = 0.005$ . The irony items score was significantly lower than that of indirect refusals ( $Z = -2.49$ ,  $p = 0.013$ ,  $d = 1.45$ ) and indirect opinions ( $Z = -2.49$ ,  $p = 0.013$ ,  $d = 1.13$ ), both with a large effect size. There was no significant difference between indirect refusals and indirect opinions ( $Z = 0.000$ ,  $p = 1.000$ ,  $d = 0.47$ ).

They also found a similar pattern in response times (p. 689):

There was a significant difference across the three item types, chi-square = 11.14,  $p = 0.004$  (Friedman test). Response times for indirect opinion items were significantly faster than for indirect refusals ( $Z = -2.76$ ,  $p = 0.006$ ), and for irony items ( $Z = -3.25$ ,  $p = 0.001$ ) (Wilcoxon signed-ranks test). Effect size was large for both,  $d = 0.79$  for the former and  $d = 0.83$  for the latter. There was no significant difference between indirect refusals and irony items ( $Z = -1.41$ ,  $p = 0.16$ ,  $d = 0.15$ ).

It should be noted here that the benchmarks used by Taguchi et al. to interpret effect size were different than what we used. Where Taguchi et al. interpreted their results as reflecting a large effect size, the Plonsky and Oswald's (2014) guidelines,<sup>47</sup> followed in the current dissertation,

---

<sup>45</sup> as well as for indirect opinions

<sup>46</sup> We will however return to an analysis of the native speaker results as they compare to the learner results in [section 7.3.4](#).

<sup>47</sup> As already stated, Plonsky and Oswald (2014) indicate that, for within-subjects comparison, for  $d$ , a small effect size lies between 0.6 and 0.99, a medium effect size lies between 1.0 and 1.39, and a large size is anything beyond 1.40. For  $r$ , an effect size of 0.25 is considered small, of 0.4 medium, and of 0.6 large

would identify only a small (or medium) effect size (cf. [sub-section 4.4.1](#)). However, the differences Taguchi et al. found in their native speakers' accuracy and comprehension speed were not present in our datasets.

In terms of accuracy, Taguchi et al.'s results reported that their native speakers were less accurate on irony items. We did not find any accuracy differences in our native speaker dataset. We know that irony can sometimes be challenging to understand, even for native speakers (Attardo, 2002; Brône; 2008; Bryant & Fox Tree, 2002; Gibbs & O'Brien, 1991; Gibbs, O'Brien, & Doolittle, 1995; Muecke, 1978; Wilson, 2012; Yus, 2000 etc. —cf. [section 2.6](#)) and that the presence of irony markers can facilitate its detection and interpretation (Muecke, 1978; Yus, 2000, etc.—cf. [sub-section 2.6.4](#)). The amount and salience of irony markers in our multimodal instrument was not a factor we could control: as a result, it is difficult to know whether Taguchi et al.'s videos contained the same quantity of contextual cues/irony markers as our videos. It is therefore possible that the native speakers in our pilot study were able to interpret the irony items more accurately because of the greater presence or salience of these irony markers. Additionally, research (Akimoto, Sugiura, Yomogida, Miyauchi, Miyazawa, & Kawashima, 2014; Channon et al. 2007; Dews, Winner, Kaplan, Rosenblatt, Hunt, Lim, McGovern, Qualter, & Smarsh, 1996; Gibbs, & O'Brien, 1991; Kuriscak, 2010, Taguchi, 2012) shows that some people are naturally better at detecting and interpreting irony than others. Given that individual skills may lead to different success rate, it is conceivable that our native speaker participants were simply better at detecting irony compared to Taguchi et al.'s group.

In terms of comprehension speed, Taguchi et al. found, on average, significantly shorter response times for indirect opinions than for irony and indirect refusals. We also observed a significant difference between indirect opinions (shorter times) and indirect refusals (longer times) but no difference with irony items. We already mentioned a potential reason for the fact that indirect refusals may have taken longer for our native speaker participants (absence of IFIDs). Additionally, and in line with what we stated above regarding accuracy, it may be the case that Taguchi et al.'s participants—and ours—were faster with certain categories (in particular indirect opinions) due the presence of salient paralinguistic cues for indirect opinions. Taguchi et al. actually credit the reliance on visual cues as the reason for faster comprehension speed in L2 learners—it is possible the same was true for native speakers. These cues may not have been as present or salient in the indirect refusal videos (in both our pilot study and Taguchi et al.'s) and irony videos (for Taguchi et al.). While irony can be challenging to comprehend

(even for native speakers—which was the case for Taguchi et al.’s participants), it seems that our native speaker participants did not encounter difficulties with the irony items they had to interpret.

Now turning our attention to the second goal of our pilot-study, we will analyze our native speaker data with the objective of identifying those items that presented some challenges for our participants so they can be eliminated from the L2 learner test.

### 6.3. Results and selection of items to be removed

Following Taguchi et al.’s protocol, the pilot study was used to refine the instrument and reduce the number of experimental items for each category of indirect meaning from 16 to 12. This section presents results from the native speaker group insofar as these results will be used to identify items to be removed in order to finalize the test that will be given to the L2 learners.

#### 6.3.1. Data analysis

Similar procedures as the previous data analysis were followed, but this time, we calculated the means per items (instead of participants) for response time and accuracy scores in each category and loaded the data files into the statistical program R (Version 4.0.3, 2020) and R Studio (Version 1.4.1103, 2021) to run the descriptive analysis of the item averages for both accuracy scores and response times.

In the section below, we will first present our descriptive statistics, and then we will describe the process by which we identified the items to be removed.

#### 6.3.2. Results

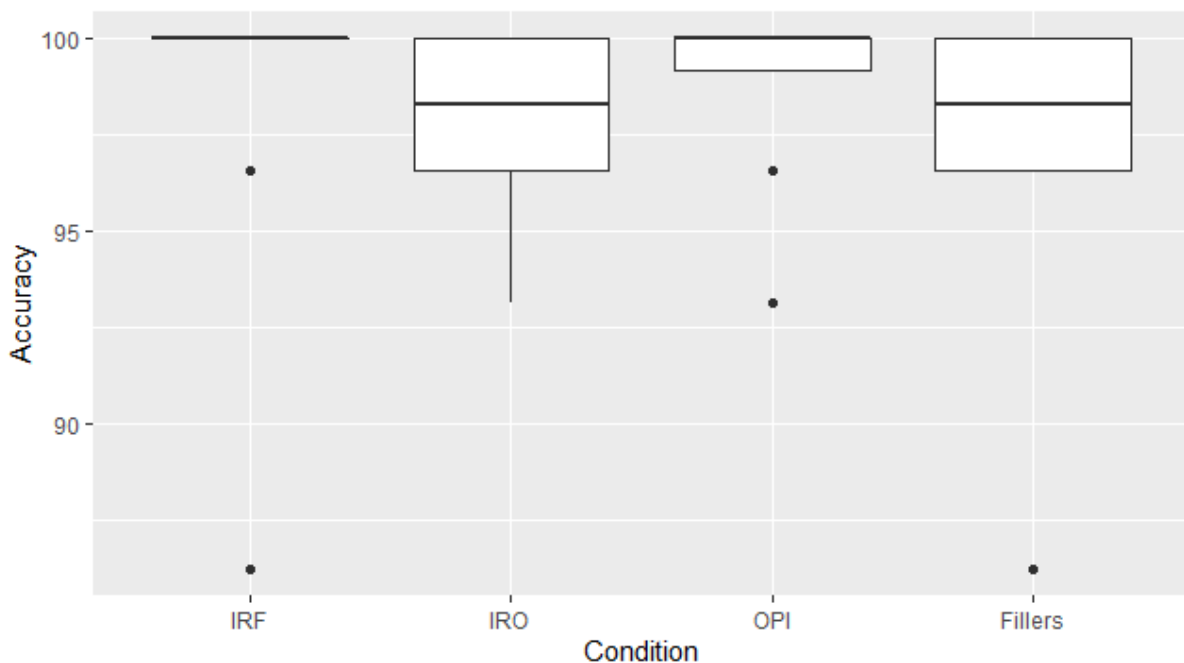
[Table 11](#) displays the results for accuracy scores (in percentage, rounded up to the 2<sup>nd</sup> decimal) per item in our native speaker dataset.

**Table 11***Item Accuracy Scores*

	Mean	<i>SD</i>	Min.	Max.
Indirect refusals	98.71	3.53	86.21	100
Irony	97.63	2.74	93.10	100
Indirect opinions	98.71	2.48	93.10	100
Fillers/Literal	96.21	5.50	86.21	100

Knowing that our native speakers performed extremely well (97.81% overall accuracy), we anticipated that identification of items to be removed based on accuracy scores may be challenging.

Using packages `dplyr`, `ggplot2` and `ggpubr` we created a boxplot of the accuracy scores per category. [Figure 11](#) displays the grouped plots with outliers (represented as dots).

**Figure 11***Items Accuracy Scores – grouped boxplot*

[Figure 11](#) confirmed a very high accuracy rate on all items, with a few outliers, which we identified by going over the spreadsheet and finding the lowest scores in each category.

Below, we will present the descriptive statistics for the native speaker response time (in seconds) for reference. We should note here that the range between minimums and maximums in [Table 12](#) seems to indicate that some questions took longer on average (sometimes close to double the amount of time), which could help with our process of item elimination.

**Table 12**

*Items Response Times*

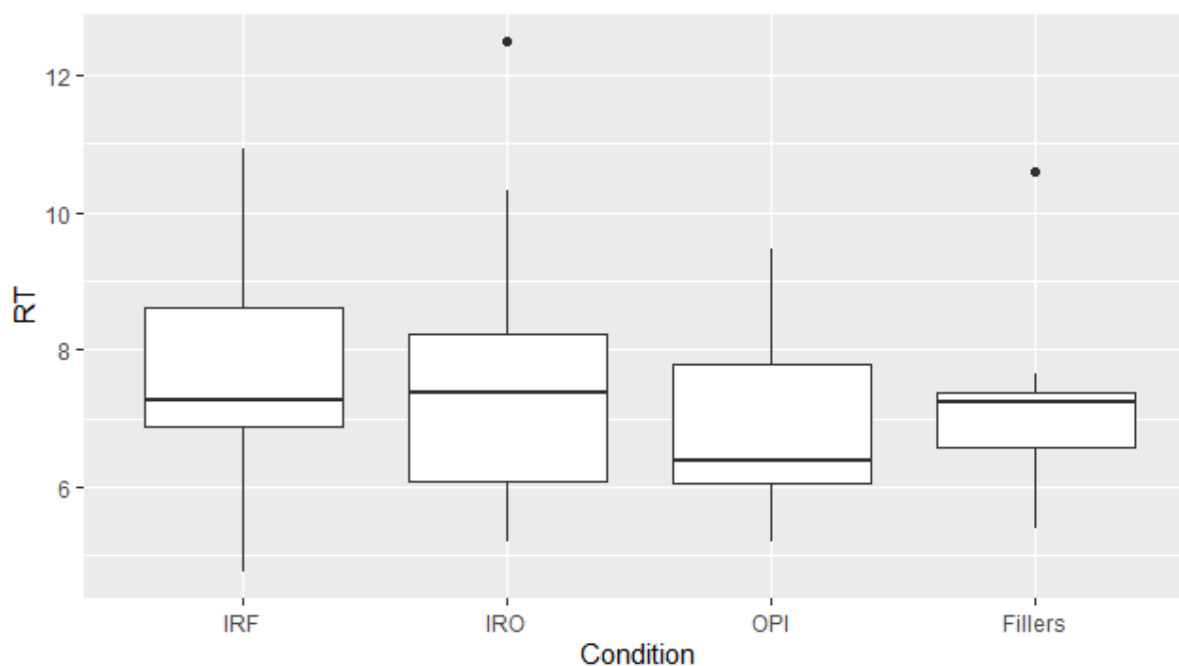
	Mean	<i>SD</i>	Min.	Max.
Indirect refusals	7.64	1.57	4.76	10.94
Irony	7.63	1.95	5.20	12.50
Indirect opinions	6.96	1.36	5.21	9.48
Fillers/Literal	7.22	1.40	5.39	10.59

Just like for the accuracy scores for each item, we created a boxplot of response times (in seconds--grouped into the four categories) using packages `dplyr`, `ggplot2` and `ggpubr` in R Studio. [Figure 12](#) displays the grouped plots with outliers (represented as dots).



**Figure 12**

*Items Response Times – grouped boxplot*



[Figure 12](#) shows a few outliers but an overall fairly consistent range of comprehension speed.

Now we will direct our attention to the scores and response times of specific items for each category in an effort to identify the items that will be excluded from the final version of our experimental study.

### *6.3.3. Identification of items to be removed*

Taguchi et al. eliminated four items per indirect meaning categories by “removing the items that had lower accuracy rates and/or longer response times than average” (p. 687). It is unclear whether they gave more weight to accuracy scores or response times, but for the most part, we opted to focus on the accuracy rate first, before then comparing it with the response time or use a combination of both in some instances. For each category, we identified the items with the lowest accuracy score and highest (i.e., slowest) response time in the native speaker pilot data, and proceeded with our elimination process as presented below (items highlighted in red below are the ones that were removed):

❖ Indirect refusal (IRF) category:

- **Accuracy:** Only three items did not receive an average accuracy score of 1 (100% accuracy) and were eliminated from the L2 test:
  - **IRF 10.** Score = 0.862, four out of 29 participants selected an incorrect answer.
  - **IRF 11.** Score = 0.966, one out of 29 participants selected an incorrect answer.
  - **IRF 13.** Score = 0.966, one out of 29 participants selected an incorrect answer.
- **Response time:** We then focused on response time. The average response time for this category was 7.643 seconds. **IRF 10** and **13** had the two highest response time (**IRF 10**, response time = 9.879 and **IRF 13**, response time = 10.936) but were already selected due to their lower accuracy. Due to a mistake, the item with the fourth slowest average response time (**IRF16**, response time = 9.016) was excluded, instead of the item with the third slowest average response time (**IRF6**, response time = 9.401). The response times for these two items were very close and we did not anticipate this error having an impact on the results.

❖ Irony (IRO) category:

- **Accuracy:** Eight items out of 12 did not receive a score of 1 (100% accuracy) in this category, ranging from 93 to 96 % (one or two participant(s) selected an incorrect answer): **IRO1**, **2**, **3**, **4**, **6**, **8**, **14** and **15**. There was not a clear difference between these scores so we decided to look at the response times, which seemed to present more significant variation, to see if we could narrow down our selection.
- **Response time:** The four items with the highest response times were:
  - **IRO 3** (response time = 8.161)
  - **IRO 6** (response time =10.317)
  - **IRO 12** (response time =12.50)
  - **IRO 15** (response time =9.951)

Since **IRO 3**, **6** and **15** were also in our list of items that did not receive 100% in accuracy, we selected them. While participants did take longer to respond to **IRO 12**, it received a score of 1 (100% accuracy) which would indicate that participants were all able to understand that item. We decided to select item **IRO 2** as our fourth item to remove in this category as it had the fourth highest response time (response time = 7.82) among the answers that had not received 100% accuracy.

❖ Indirect opinion (OPI) category:

- **Accuracy:** Four items out of 12 did not receive a score of 1 (100% accuracy) in this category, ranging from 93 to 96 % (One or two participant(s) selected an incorrect answer): **OPI 3, 11, 13 and 15**. These four items were automatically selected for removal but we did double check response time for consistency.
- **Response-time:** The average response time for this category was 6.954 seconds.
  - **OPI 11 and 13** had the two highest response times (respectively 9.421 and 9.483) in this category.
  - Response times for **OPI 3 and 13** were also above average (respectively 7.164 and 7.561)

Since our pilot study only had 10 “filler” items and Taguchi et al. did not eliminate any of them, we kept the same ones for our L2 test. In the end the following items were removed for the L2 study

- **IRF 10, 11, 13 and 16**
- **IRO 2, 3, 6 and 15**
- **OPI 3, 11, 13 and 15.**

We decided to use two of the discarded items as practice items for the L2 test. We selected IRF 11 (score = 0.96, response time = 7.73) and OPI 3 (score = 0.96, response time = 7.16), based on the fact that their accuracy score was still very high and their average response time close to the average response time for their respective category.

## Chapter 7 - Experimental study – L2 learners

Our goal for this chapter is to focus on the quantitative portion of our experimental study with the L2 learners. First, we will describe the methods used for the administration of our learner survey as well as the data analysis ([section 7.1](#)). Following the Methods section, we will present the results of the quantitative analysis performed on the learner datasets ([section 7.2](#)), which aim to answer the following research questions:

- **To what extent are L2 French learners able to understand indirect meaning?**
- **How do different types of indirect meaning impact L2 French learners' comprehension?**

Our discussion ([section 7.3](#)) will be organized in several sections: first we will review our results and answer our research questions ([sub-section 7.3.1](#)). Then we will interpret our results with regard to conventionality ([sub-section 7.3.2](#)). Finally, we will compare our findings with Taguchi et al. ([sub-section 7.3.3](#)). To that effect we will also reprocess our original native speaker dataset<sup>48</sup> to only include items that were used in the L2 learner analysis ([sub-section 7.3.4](#)).

### 7.1. Methods

In this section, we will first describe our participants ([7.1.1](#)) and the outcome of their proficiency self-assessment survey ([7.1.2](#)). We will then turn our attention to the test administration procedures ([7.1.3](#)) and finally present our data analysis protocol. Results and discussion will be presented in sections [2](#), and [3](#) of the present chapter.

#### *7.1.1. Participants in the L2 study*

Students were recruited from two undergraduate advanced-level French courses at a public university in the United States. Due to course and institutional constraints, we worked with intact classes only, but students had the option to opt out of the study (see IRB requirements in [sub-section 7.1.3.1](#)). In total, 37 students completed the data-collection activity and all agreed

---

<sup>48</sup> Our previous native speaker analysis ([section 6.2](#)) was based on all original items ( $n = 58$ ), but in order to directly compare the native speakers and L2 learners in our study, you re-ran our analysis using the 48 items used for the L2 learner test.

for their answers to be used for this study. Four students were self-reported native speakers of French, and their answers were excluded from the results. Our total number of participants was 33: 26 females, six males, and one other, with an average age of 22 years old. All but one participant selected for the study indicated they were native speakers of English. The one participant who indicated her first language to be Spanish had been living in the US for 19 years (since the age of 3) and can be considered bilingual. Thirteen of the 33 participants had just finished the Intermediate level French course and were in their fifth semester of French. The rest of the participants ( $n = 20$ ) had already taken advanced-level French courses prior to that semester (i.e., this was not their first semester taking advanced French, which meant they would range from sixth to eighth semester). Seven participants reported having studied or lived in a French-speaking country but upon request for clarification, none of them had stayed in the country for more than one semester.

Taguchi et al. recruited students enrolled in intermediate and advanced level classes at a private institution (ranging from third to sixth semester of language study). To follow that model, we had initially planned on recruiting additional participants from a third course at the intermediate level. However, after looking at the preliminary results from our self-assessment survey (which will be discussed in the [next sub-section](#)), it became evident that Intermediate-level students at the researcher's institution would not be proficient enough (as compared to Taguchi et al.'s participants).<sup>49</sup> Additionally, the two advanced-level French courses that we selected for this project already represented different institutional levels since the courses were available to students in their fifth (right after the Intermediate level course which would be equivalent to Junior level), all the way to eighth (or more) semester (graduating Senior level). We hypothesize that the proficiency level between students who had only completed four semesters of language studies and that of students who had already completed one or more advanced courses (five semesters and more) would be different. A closer look at the results of the self-assessment survey confirmed this assumption.

### *7.1.2. Learners' self-assessment of language proficiency*

Following Taguchi et al.'s protocol, we used the NCSSFL-ACTFL Can-Do self-assessment

---

<sup>49</sup> As mentioned in Chapter 4, [sub-section 4.3.1](#), language students in a third semester course of French language at an American university are usually expected to be at an Intermediate Low level on the ACTFL scale, which corresponds to level A1.2/A2.1 of the CEFR.

checklist to assess the proficiency levels of the students in the two advanced-level French courses selected for the study. The self-assessment survey we designed used 15 items from the Intermediate/Intermediate-High and Advanced-Low levels<sup>50</sup> of the ACTFL Can-Do self-assessment. As a point of reference, Intermediate-High usually corresponds to a rating of B1.1 on the Common European Framework of Reference for Languages (CEFR) and Advanced-Low a CEFR rating of B1.2,<sup>51</sup> which is where we would expect students at a Junior or Senior level to be at our institution. The survey was administered online about two weeks before the study. Students were asked to report their abilities using a five-point assessment scale ranging from (1) “I can’t do this/It’s too difficult” to (5) “I can do this/No problem.” The Can-Do statements concern tasks that L2 learners should be able to carry out in different modes of communication (interpersonal, interpretative and presentational). In their paper, Taguchi et al. did not explicitly indicate which statements were used and only provided some examples. Therefore, following Taguchi et al.’s selection process (p. 683 “statements that closely reflected activities and tasks in the participants’ Spanish classes”), we selected items that were in line with activities and tasks that the participants usually complete in their French courses. Sample statements included “I can read movie reviews to choose what to watch,” or “I can understand some events described in an excerpt from a historical journal.” The complete list of statements used for this study can be found in [Appendix J](#).

Thirty-six students completed the self-assessment survey, but we eliminated the answers from the four students who indicated they were native speakers of French. Out of 32 students remaining, 10 students reported being French majors, and 17 were French minors. It is important to note here that four of the students who took the self-assessment test across both classes were not present on the day we collected our data in class, and five students who participated in the L2 test did not complete the self-assessment survey despite our repeated request (two students submitted theirs later in the semester but by then we had already analyzed the self-assessment data to evaluate the general level in the classes). We did not think this would pose a problem since the main purpose of this self-assessment was to gauge the general proficiency level of the students in the two classes we selected. The results of the self-

---

<sup>50</sup> Intermediate Can-Do Statements: <https://www.actfl.org/sites/default/files/can-dos/Intermediate%20Can-Do%20Statements.pdf>

Advanced Can-Do Statements: [https://www.actfl.org/sites/default/files/can-dos/Advanced%20Can-Do\\_Statements.pdf](https://www.actfl.org/sites/default/files/can-dos/Advanced%20Can-Do_Statements.pdf)

<sup>51</sup> Looking at comparative rating for receptive skills (reading and listening): [https://www.actfl.org/sites/default/files/reports/Assigning\\_CEFR\\_Ratings\\_To\\_ACTFL\\_Assessments.pdf](https://www.actfl.org/sites/default/files/reports/Assigning_CEFR_Ratings_To_ACTFL_Assessments.pdf)

assessment test were fairly consistent with the ACTFL intermediate to low advanced level of proficiency, with an average rating on a five-point scale was 3.72 ( $SD = 0.54$ , range = 2.73-4.76), for the 15 items that were taken directly from the intermediate/low advanced statements.

It should be noted that, based on this self-assessment, our participants' reported proficiency seemed lower than what Taguchi et al. reported (their average was 4.24 across the intermediate and advanced courses) but this was another argument in favor of not collecting data in our Intermediate-level classes.<sup>52</sup> Although self-assessment of language proficiency has been questioned, in the institutional context in which the current experiment was conducted it was the only feasible means for assessing language proficiency. Interestingly, only half of the students who were in their fifth semester actually rated themselves lower than average and nine out of eleven students who were in their sixth or more semester rated themselves below average (which could be a sign that the self-assessed proficiency level of the students was not necessarily consistent with their institutional level, or a reflection of the lower level of confidence in actually more proficient learners).

### 7.1.3. Procedures

We will now turn our attention to the procedures of the experiment. The next sections will focus on the Institutional Review Board approval process and the survey administration.

#### 7.1.3.1. IRB

The American university where we completed this project has a protocol in place to comply with federal regulations, and all researchers must first seek approval from their institution's Institutional Review Board (IRB). At time of testing in France, similar regulations were being discussed at the French university with which the researcher is associated, but had not yet been put into place.

---

<sup>52</sup> From a practical point of view, the Covid-19 health crisis limited our data collection timeframe and ability to work with additional classes. However, while we could have moved forward with data collection in an Intermediate course, it would have been unnecessary from a replication standpoint, given that students in our two advanced-level courses already self-reported their proficiency level to be intermediate/low advanced based on the ACTFL scale. As discussed in sub-section 4.3.1, it is surprising that students in a third semester language course could actually reach an intermediate-high/advanced low level as defined by ACTFL, which is what was reported in Taguchi et al. Without knowing which Can-Do statements Taguchi et al. used, it is impossible to exactly compare their participants' self-assessed proficiency with our participants' level.

Our Informed Consent Form (ICF—see [Appendix K](#)) was to be delivered electronically at the beginning of the computer survey that our L2 learners would complete . While all participants (captive audience of students enrolled in a specific course/section) had to complete the activity used to collect data as part of a regularly-scheduled assignment for that course, the consent form gave them the option to agree (or not) for their data to be used for our study. Additionally, they had the option to indicate whether they did not wish to be contacted for a follow-up interview, if selected to be participants in the retrospective interviews.

#### 7.1.3.2. Survey administration

The survey was administered during an in-class activity done in two different advanced-level French courses, meeting for 90 minutes. We anticipated that it would give us enough time to provide students explanations and clarifications, and to handle any technical issues we might have encountered before students started the activity, with sufficient time remaining for them to finish it without having to rush through the tasks.

Two weeks prior to the activity, we created a course on Canvas (our Learning Management System) and invited all the students in both classes to join the online course so they could have access to a copy of the consent form, in the event that they wanted to email the researcher with questions (following IRB protocol). The Canvas course was also used to send students reminders and share the link to the self-assessment test, which they completed outside of class, prior to the in-class administration of the online test. On the day of the activity, students met the researcher in a computer lab on campus. Each participant sat at a table and used a PC and a set of headphones.

At the beginning of class, the researcher briefly went over the consent agreement again, and students accessed the online survey through a link that was posted in the Canvas course. The researcher read the instructions (see [Appendix H](#)) to the students out loud as they were looking at them on their screen, and then they were instructed to wait for the researcher’s signal before moving on to the next page. After asking participants to check the volume in their headphones (one set was not working and the participant had to move to a different computer), participants were instructed to click “next” and start playing the first video, which was one of the two practice items, and the same for all participants. Once all participants watched the video, they



waited for our signal to click on “next” to access the multiple-choice answer page. We reminded them that the timer would start as soon as the MCQ would load on their screen and would stop once they selected their answer. We also explained that they could not change their answer once selected as responding to the MCQ would automatically advance the survey to the next page. Participants then proceeded to the next page and the second practice item.

During this “practice” phase with our first group of participants (first class), it became evident that the videos were still taking longer than expected to load. Participants would start watching a video, and it would buffer and stop. They would then have to wait or sometimes restart the video. Because we did not want the participants to watch the video multiple times (and they were also reminded of this requirement), we instructed them to wait a few seconds every time they landed on the video page to give time for that video clip to load. We recommended they wait until the video was loaded up to two-thirds (as seen on the video loading bar that appeared at the bottom of the player), as that seemed to solve the buffering issues. This solution was adopted by the participants who were then able to watch all the videos without any major problems. Waiting a few seconds before playing each video did not impact the results of our study, since response time was only measured on the MCQ page. In the end, this may have only added a few minutes to the total activity time, but it also did not seem to pose any problems since all the participants were able to complete the survey before the end of class time. Similar instructions were provided to the second group of participants (second class). At the end of the two practice items, we asked participants if they had any questions or concerns, and then told them to continue with the activity. The 46 items of the actual experiment were presented in a randomized order for each participant.

The rest of the survey administration went smoothly. We did not run into additional technical issues and participants finished the activity in 41 minutes on average (from the moment they opened the survey and we went over the consent form, to the time they closed the survey after answering the last question. Range = 28.5-50 min,  $SD = 5$ ) across both classes. During the activity, the researcher walked around the room to monitor participants’ advancement, or spot any potential issues. All participants appeared to follow the provided instructions and watch each video one time only. This observation was later confirmed by the results of our video viewing confirmation question at the end of the survey: When asked “Did you only watch each video once?,” 31 out of 33 students answered “yes”—which is consistent with what we had witnessed during the activity—and the two participants who answered negatively indicated that

they estimated having watched the videos multiple times only 5-10% of the time/of the videos, but it is not clear whether they included the two practice items in that percentage, since we know students had buffering issues at the beginning which caused them to have to restart the video clips a couple of times. Additionally, it is possible that when they did watch a video multiple times, they could have only replayed a small portion of the video (for example the beginning of the video when experiencing buffering or the end of a video).

#### *7.1.4. Data analysis*

Our first and second research questions aimed to investigate possible difficulties and differences in L2 French learners' comprehension accuracy and response speeds depending on different types of indirect meaning. As a reminder, accuracy and speed were operationalized in the following ways:

- Comprehension accuracy: percentage score<sup>53</sup> for each type of indirect/literal meaning. A score of 1 was given to all correct answers and a score of 0 to incorrect answers, following Taguchi et al.'s system.
- Comprehension speed: response times based on the average number of seconds it took participants to answer questions correctly in each category.

Just like with the pilot study, we exported the Qualtrics results into an excel sheet and then separated the data into two sets: one for the accuracy results and one for the response times. We removed the incorrect answer response times from our dataset (following Taguchi et al.) so that we could look at potential differences in response times when the participants had successfully understood the intended meaning of the utterances. We explored normality using the `pastecs` package in R Studio (Version 1.4.1103, 2021) with R (Version 4.0.3, 2020). Unlike Taguchi et al.'s datasets, our learner data were normally distributed: using the `stat.desc` command, we looked at the kurtosis and skewness of all four item types for both accuracy and response times and we did not find a significant skew or kurtosis, which means we could assume a normal distribution. As a result, we were able to use parametric tests to analyze our two learner datasets.

To answer our first two research questions, we used a multilevel linear model following Field et al. (2012, p. 573) to analyze response accuracy and speed. For this, we used the `nlme` package

---

<sup>53</sup> As previously mentioned, Taguchi et al. used a 0 to 12 interval scale for accuracy but since our literal meaning category only included 10 items, we had to convert the accuracy scores into percentages, in order to include the literal meaning category in our statistical analysis later on.

in R Studio (2021). This type of model was selected because it allowed us to account for the fact that each participant was exposed to each of the four conditions (indirect refusal, irony, indirect opinion, and filler, i.e., literal meaning). More specifically, we included a random intercept for each Participant, as well as random slopes for the Participant\*Item type interaction. The continuous dependent variable for the first dataset was accuracy (in percentage), and for the second response time. In addition to the random effect structure, we also included the fixed effect of Item type in the model, in order to evaluate its potential impact on accuracy scores and response speeds. Filler was used as the reference category for this factor. In order to determine whether Item type significantly impacted accuracy scores and response times, we conducted an ANOVA comparison between the baseline model (i.e., the model containing only the random effect structure) and the model containing the random effect structure and the fixed effect. If the model containing the fixed effect (i.e., item type) were to be significantly better than the baseline model, then we can conclude that the inclusion of Item type significantly improved the model. Finally, we computed post hoc comparisons using Tukey HSD (with the multcomp package) in order to determine which item types significantly differed from each other. Here again, effect sizes (both  $d$  and  $r$ ) were calculated for each comparison using the compute.es package in R Studio (2021), and were interpreted by following Plonsky and Oswald's (2014) recommendations. For within-subjects comparison, for  $d$ , a small effect size lies between 0.6 and 0.99, a medium effect size lies between 1.0 and 1.39, and a large size is anything beyond 1.40. For  $r$ , an effect size of 0.25 is considered small, of 0.4 medium, and of 0.6 large.

## 7.2. Results

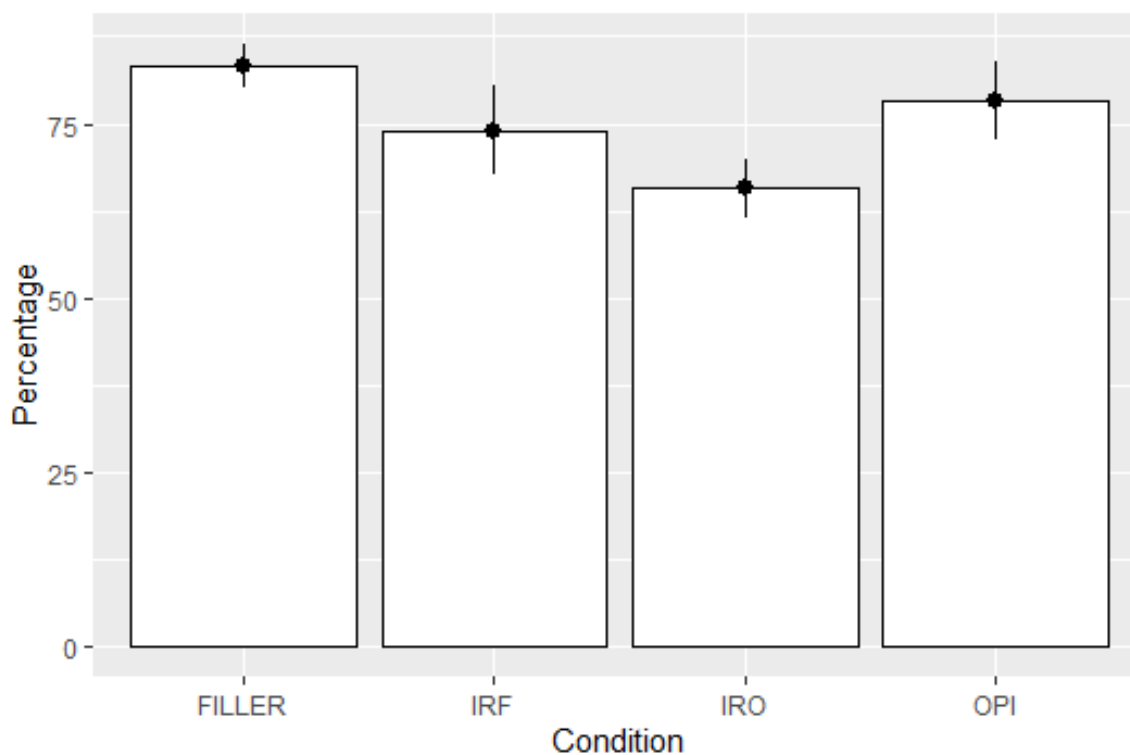
First, we will provide results for comprehension accuracy and then for comprehension speed.

### 7.2.1. *Comprehension accuracy*

[Table 13](#) below displays descriptive statistics for the L2 French learners' comprehension accuracy scores (in percentage, rounded up to the 2<sup>nd</sup> decimal). [Figure 13](#) presents a bar chart of these scores for each category/condition.

**Table 13***Learners' Accuracy Scores*

Item type	Mean	SD	Min.	Max.
Indirect refusals	73.99	17.89	33.33	100
Irony	65.66	12.28	41.67	91.67
Indirect opinions	78.28	16.13	41.67	100
Fillers/Literal	83.33	8.90	70.00	100
Overall	75.32	13.80	46.67	97.92

**Figure 13***Learners' Accuracy Scores - bar graph with error bars*

The above graphic seems to indicate that there are some differences in comprehension accuracy between the categories. Indeed, learners' scores were lower for irony but higher for the literal meaning items (fillers). This was confirmed with statistical analysis.

Our multilevel linear model showed that the variable of Item type had a significant impact on the accuracy of responses. The details for this final model are provided in [Table 14](#).

**Table 14***Accuracy Model results*

Factor	Estimate	SE	t value	Confidence intervals	
				Lower	Upper
(Intercept)	83.333	2.478	33.634	78.490	88.176
Item type [Filler]					
Indirect opinion	-5.051	2.596	-1.946	-10.125	0.024
Indirect refusal	-9.343	2.596	-3.599	-14.418	-4.269
Irony	-17.677	2.596	-6.809	-22.751	12.603

Given that the estimates for our three categories of indirect meanings in [Table 14](#) have negative estimates, we know--in absolute terms--that indirect opinion, indirect refusal and irony items led to less accurate responses as compared to our literal meaning items (fillers). Details for the random effect of participants are presented in [Appendix L](#).

Post hoc comparison tests helped us determined which of the comparisons were significant. The Tukey HSD test indicated that the mean accuracy scores on the following comparisons were significantly different: IRO vs. OPI ( $p < 0.001$ ), IRO vs. FILLER ( $p < 0.001$ ), IRO vs. IRF ( $p = 0.00599$ ) and IRF vs. FILLER ( $p = 0.00139$ ). The OPI vs. IRF ( $p = 0.33458$ ) and the OPI vs. FILLER ( $p = 0.19700$ ) comparisons did not reveal significant differences. The effect size for the comparison between IRO and FILLER was large ( $d = -1.65$ ,  $r = -.64$ ), whereas the comparison between IRO vs. OPI showed a small to medium effect size ( $d = -0.88$ ,  $r = -.41$ ) and the IRO vs. IRF comparison showed a borderline small effect size ( $d = -0.54$ ,  $r = -0.27$ ). As for the comparison between IRF and FILLER, the effect size calculations revealed a small effect ( $d = -0.66$ ,  $r = -0.32$ ). While the effect size varies between the different pairs, irony stands out as the condition that may have impacted comprehension the most.

### 7.2.2. *Comprehension speed*

[Table 15](#) displays descriptive statistics for the learners' comprehension speed (in seconds) on correctly answered questions.

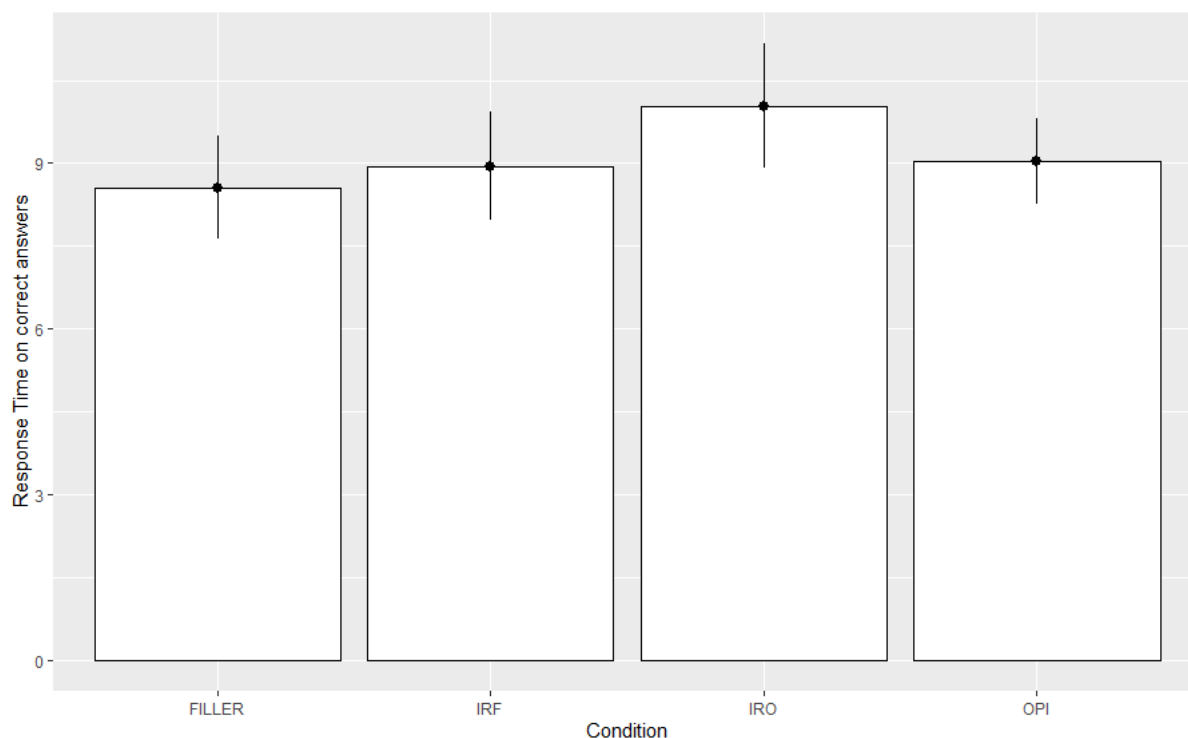
**Table 15***Learners' Response Times on correct answers*

Item type	Mean	<i>SD</i>	Min.	Max.
Indirect refusals	8.37	2.53	4.22	14.21
Irony	9.44	3.11	4.09	20.31
Indirect opinions	8.77	2.10	4.52	13.33
Fillers/Literal	7.86	2.32	3.31	13.25
Overall	8.61	2.51	4.04	15.28

We can see that, among the correct answers, the category that took participants the longest was irony. The literal items were answered correctly the fastest. As a point of comparison, we also looked at the average response time for the full dataset (including both correct and incorrect answers) and the response times on correctly answered items was lower (i.e., faster) than the overall response time for the full dataset (8.61 vs. 9.15 sec). This seems to indicate that, on average, participants took longer to answer questions that resulted in an incorrect answer. It is also worth noting that once we removed the incorrect answers from our data, the standard deviation for the means in each category also went down (2.51 vs. 2.69). A table of the learners' response times showing the descriptive statistics for the full dataset is included in [Appendix M](#). [Figure 14](#) presents a bar graph of these response times for each condition based on correct answers.

**Figure 14**

*Learners' Response Times on correct answers - bar graph with error bars*



The chart seems to indicate that there are some differences in comprehension speed between the item types. Indeed, learners' response times were slower for irony but faster for the literal meaning items (fillers). We next ran a multilevel learner model to test for significance.

Our multilevel linear model showed that the variable of Item type had a significant impact on response time. The details for this final model are provided in [Table 16](#).

**Table 16**

*Response Time Model results*

Factor	Estimate	SE	t value	Confidence intervals	
				lower	upper
(Intercept)	8.562	0.473	18.110	7.637	9.486
Condition [Filler]					
Indirect opinion	0.476	0.325	1.465	-0.159	1.112
Indirect refusal	0.388	0.325	1.192	-0.248	1.023
Irony	1.478	0.325	4.545	0.842	2.114

[Table 16](#) shows us that—in absolute terms—the questions associated with indirect meaning items (indirect opinions, indirect refusals and irony) took longer to answer than those associated with the literal meaning items (fillers). Our random effects are presented in [Appendix N](#).

Post hoc comparisons using the Tukey HSD test indicated that the response times on the following comparisons were significantly different: IRO vs. FILLER ( $p < 0.001$ ), IRO vs. IRF ( $p = 0.00350$ ) and IRO vs. OPI ( $p = 0.00961$ ). The IRF vs. FILLER ( $p = 0.61976$ ), OPI vs. IRF ( $p = 0.99260$ ) and the OPI vs. FILLER ( $p = 0.444484$ ) comparisons did not reveal significant differences. However, the effect size for the three pairs that were significantly different were all below the thresholds for “small” suggested by Plonsky and Oswald (2014): IRO vs. FILLER ( $d = 0.4$ ,  $r = 0.2$ ), IRO vs. IRF ( $d = 0.21$ ,  $r = 0.1$ ) and IRO vs. OPI ( $d = -0.11$ ,  $r = -0.05$ ). Therefore, we can say that while there seemed to have been some differences in response times between irony and the other categories, the size of the effect was negligible.

### 7.3. Discussion

The analysis of our quantitative data allowed us to respond to our first two research questions: **To what extent are L2 French learners able to understand indirect meaning? Do different types of indirect meaning impact L2 French learners’ comprehension accuracy and speed?** Answers to these questions will be presented in [sub-section 7.3.1](#). In [sub-section 7.3.2](#), we will interpret these results with regard to the notion of conventionality. Finally, in [sub-section 7.3.3](#), we will compare our findings with that of Taguchi et al.’s.

#### *7.3.1. Answers to our research questions*

With an overall accuracy of 75.32%, we know that our L2 learners’ comprehension was somewhat challenged, and it seems that indirect meaning was generally harder to comprehend since the literal meaning category resulted in the highest accuracy rate (83.33%). Together, all three categories of indirect meaning only reached an average accuracy rate of 72.64% (cf. [sub-section 7.2.1](#)). However, to answer our first two research questions, we need to look at the different categories of indirect meaning separately because not all of them resulted in significant



differences compared to literal meaning. That means that some types of indirect meaning were not actually harder to understand than literal meaning.

First, we noted a more significant impact of irony, which resulted in lower accuracy scores and slower comprehension speed, both of which reached statistical significance in comparison with all three other item types (indirect opinions, indirect refusals and literal items). This suggests that our L2 learner group had more difficulties understanding irony items than they did the other indirect or literal meaning categories. The fact that our L2 learners struggled with irony is consistent with previous research findings (Bouton, 1992, 1994b, 1999; 1999; Bromberek-Dyzman & Rataj, 2016; Kim, 2013, 2014, 2016; Shively, 2013; Shively et al., 2008; Taguchi et al. 2016). Irony is a complex phenomenon, and we know that even native speakers can sometimes have trouble understanding ironic comments. That is because irony processing relies on the hearer (or observer)'s ability to detect an incongruity or incompatibility between what is said and what is implied based on the context or situation. Therefore, access to irony interpretation often requires processing of contextual cues (cf. Attardo 2000b, Attardo & Poggi, 2003, Attardo et al., 2013, Yus, 2000—cf. [section 2.6](#)) as well as processing of the linguistically-encoded meaning of the ironic comment (literal meaning) through a combination of top-down and bottom-up strategies. In [Part III](#) of this dissertation, we will investigate the use of these strategies by a selected group of our L2 learners to find out whether this complex process of inferencing is what caused our participants some difficulties.

While we found that irony was more challenging to understand than literal items (statistical difference of a large effect size), we did not observe any statistical differences between indirect opinions and literal items (for both accuracy and speed) and the effect size for the difference in accuracy between indirect refusals and literal items (the only significant difference we found besides irony) was small. This seems to indicate that the indirect nature of the two indirect speech acts we studied only minimally impacted comprehension.

Our L2 learners did not show significant differences in processing times among indirect opinions, indirect refusals or literal meaning items. The only difference we found in comprehension speed for our L2 learner group was between irony and other categories, which could be evidence of a link between comprehension difficulties and processing speed (since participants were also less accurate on this category). However, the effect size for the

differences in irony comprehension speed was extremely small, which indicates that the different categories of meanings did not have a strong impact on response time overall.

What do our results mean with regard to Taguchi's claim that conventionality of indirect meaning impacts comprehension? To answer this question, we will now interpret the above findings based on whether the indirect meaning was conventional or not.

### 7.3.2. *Conventionality effect*

In her previous studies,<sup>54</sup> Taguchi (2002, 2005, 2007, 2008a, 2008b, 2008c, 2009a, 2009b, 2011; and Taguchi et al., 2013) had consistently observed that indirect refusals (which she considered conventional) were easier for L2 learners (in L2 English, L2 Japanese and L2 Chinese) to understand than indirect opinions (which she considered nonconventional). Based on Taguchi's hypothesis that conventionality facilitates comprehension, we should expect indirect refusals and irony to be easier to understand than indirect opinions. This was not the case in our study: there was no significant difference between indirect refusals and indirect opinions, and irony was significantly harder to understand than indirect opinions. If conventionality was a factor, we would have expected our L2 learners to struggle with non-conventional indirect opinions more than with conventional indirect refusals (as was the case in all of Taguchi's previous studies).

One issue to consider for this discussion on the impact of conventionality is the problematic categorization of our three categories of indirect meaning by Taguchi et al., as previously mentioned. Indeed, they classified indirect opinions as nonconventional, but we identified potential conventional patterns in indirect opinions (adjacency pair: asking for someone to provide opinion, which sets an expectation for the answer to include some kind of evaluative judgement—cf. [section 2.7](#)). In our test, nine out of the 12 indirect opinion dialogues followed this type of adjacency-pair pattern as shown in the following example:

**OPI 16** - Conversation entre deux amis, Julien (homme) et Juliette (femme). Ils discutent de leurs préférences en matière de littérature.

**Julien:** J'ai enfin pu acheter le dernier livre de la série que j'adore.

**Juliette:** Tu veux dire ceux avec des zombies qui parlent de l'apocalypse?

**Julien:** Oui! J'aurai fini en deux jours! Est-ce que tu aimes ce genre de romans?

**Juliette:** Je préfère les histoires romantiques ou les livres qui parlent de voyage.

---

<sup>54</sup> See [Appendix A](#) for a summary of Taguchi's studies.

The lack of conventionality claim that Taguchi made regarding indirect opinions may not therefore be valid. On the other end of that argument, we think it is possible that the challenge our L2 learners faced with irony could lie in the fact that most of our dialogues containing irony items did not follow a predictable pattern of interaction (such as the adjacency pair rule) as opposed to our indirect refusals or some of our indirect opinions. Indeed, most ironic statements in our irony dialogues (nine out of 12) came as unsolicited opinions—cf. [sub-section 4.1.4](#)—and the fact that they did not follow the same question-answer pattern of discourse as the two indirect speech acts in our study could have contributed to the L2 learners’ difficulties in anticipating and interpreting the message conveyed by the ironic comments. Additionally, we had previously stated that meaning replacement irony did not seem to follow a very predictable pattern (cf. [sub-section 4.1.3](#)). Along the same line, we questioned the lack of IFIDs use for indirect refusals in French and English (cf. [sub-section 2.7.2](#)). We know from our native reviewers that some of our indirect refusal target sentences (with no IFID) did not feel natural to them based on the contexts of our dialogues, and we argued that this may have been the reason why our native speaker group took longer to comprehend indirect refusals in our pilot study (cf. [sub-section 6.2.3](#)). We cannot know whether that was the source of some difficulties for our L2 French learners, but this issue again challenges the categorization of indirect refusals as conventional.

Another possible factor to consider to explain the lack of conventionality effect in our study (which Taguchi et al. also mentioned in their study) is the fact that previous studies from Taguchi did not include videos. It is possible that the impact of conventionality she found in her previous studies was offset by the access to visual cues in our study: the presence of these paralinguistic cues could have facilitated the comprehension of nonconventional indirect meaning items, thus limiting the impact of conventionality. Our analysis of inferencing strategies used by some of our L2 learners may once again shed some light on this potential phenomenon.

This unexpected finding regarding the lack of conventionality impact on L2 learners’ comprehension was also found by Taguchi et al. in their 2016 paper. We will now compare our results to theirs to see whether our close replication identified other similarities or differences.

### 7.3.3. Comparison with Taguchi et al. (2016)

Converting Taguchi et al.’s accuracy scores for each category (they used a point system in their paper—cf. [Table 2, section 4.4](#)), we obtained the following accuracy percentages that can be compared with our results:

**Table 17**

*Comparison of Learners’ Accuracy Scores between Taguchi et al. and our data*

	Taguchi et al. (2016)	Our data
Item type	Mean	Mean
Indirect refusals	89.83	73.99
Irony	75.75	65.66
Indirect opinions	87.75	78.28
Overall	84.44	72.64

First, we notice that Taguchi et al.’s participants performed better than our L2 learners overall with almost 10% better accuracy (across the board and in each category). We suspected that participants in Taguchi et al.’s study had a higher proficiency level than our L2 learners (cf. [sub-section 7.1.2](#)) and this gap in overall accuracy is likely the result of a proficiency difference. Like us, Taguchi et al. found that “irony items were the most difficult to comprehend” (p. 692) based on accuracy scores (p. 689):

The Friedman test revealed a significant difference in L2 Spanish learners’ comprehension scores across three item types, chi-square = 27.66,  $p = 0.0001$ . The Wilcoxon signed-ranks test showed that the irony items score was significantly lower than that of indirect refusals ( $Z = -4.45, p = 0.0001$ ) and indirect opinions ( $Z = -3.19, p = 0.0001$ ). Based on Cohen’s  $d$ , effect size was large for the former ( $d = 1.01$ ) and medium-large for the latter ( $d = 0.76$ ). There was no significant difference between indirect refusals and indirect opinions ( $Z = -0.71, p = 0.48, d = 0.16$ ).

They determined a large effect size for the differences between irony and indirect refusals, and a medium-large effect size between irony and indirect opinions. We observed smaller effect sizes, but using Plonsky and Oswald’s (2014) recommended benchmarks, Taguchi et al. would have found that effect size to be borderline medium and small respectively, which is closer to

our findings. As such, our replication confirms Taguchi et al.'s findings with regard to irony comprehension.

Taguchi et al.'s participants performed better on indirect refusals than on the other two categories of indirect meaning (highest accuracy rate), but because the difference was not statistically significant, Taguchi et al. could not corroborate previous findings (Taguchi, 2002, 2005, 2007, 2008a, 2008b, 2008c, 2009a, 2009b, 2011) with regard to the impact of conventionality on L2 English comprehension. In our case, our L2 learners were less accurate with indirect refusals (which is conventional) than with indirect opinions (non-conventional) and literal meaning (the difference with the latter reached statistical significance), a result that seems to go against the conventionality effect since we would expect indirect refusals to be easier. The validity of this unexpected finding—which did not uphold Taguchi's previous claim that conventionality facilitates comprehension of indirect meaning—was reinforced by our replication study: comparison of indirect refusals and indirect opinions (both accuracy and speed) in our study did not show any statistically significant differences, which allow us to expand Taguchi et al.'s findings to a different language (French).

Now turning our attention to comprehension speed: while Taguchi et al. found that irony items were harder to comprehend (which is consistent with our findings), “they were not the slowest items to comprehend” (p. 692): indirect refusals were (even though they were the easiest to understand). Taguchi et al.'s participants were significantly faster in comprehending indirect opinions than irony and indirect refusals (no statistical difference in response times was found irony and indirect refusals). Taguchi et al. argue (based on the results from their retrospective verbal interviews) that indirect opinions were processed faster due to their participants relying on visual cues to access the intended meaning (cf. [section 3.2](#)). They state that (2016, p. 694):

[V]isual cues reduced the cognitive load, which in turn freed up resources to be allocated to other areas of processing (e.g., reading multiple-choice options to identify the correct option), leading to shorter response times.

Our findings with regard to comprehension speed do not align with that of Taguchi et al. In our case, irony was significantly slower than all other categories of indirect meaning. We believe that both the slow response times and low accuracy scores are due to the complexity of the irony implicatures and the need to rely on both bottom-up and top-down strategies using multiple contextual cues which are made difficult by the L2 learners' relatively low proficiency level. Because conventionality of irony is based on a pattern of meaning reversal or meaning

replacement, the necessary criteria to detect, process and interpret it is first to access the sentence-level/literal meaning of the utterance.

One of the puzzling aspects of Taguchi et al.'s study is the fact that the native speakers in their pilot study performed in much the same way as their L2 Spanish learners: in terms of accuracy, “the same pattern was found in the native speaker data,” and “[s]imilar to the accuracy results, patterns in native speakers’ response times were exactly the same” (p. 689). Both groups had issues with comprehending irony. This finding seems problematic to us since Taguchi et al. interpret their learner patterns as being specific to L2 learners. In our opinion, the fact that both native speakers and L2 learners encountered similar difficulties with certain types of indirect meaning (mainly irony), would suggest that the problem is not necessarily related to L2 comprehension, but rather a difficulty with indirect meaning (specifically irony) in general (for both native speakers and L2 learners). In order to ensure that our data did not present this issue, we decided to reprocess our native speaker dataset to only include those items that were used in the L2 learner test, and reanalyze it to see if the patterns would be different.

#### *7.3.4. Reprocessing of the native speaker dataset*

In [Chapter 6 \(sub-section 6.3.2\)](#) we discussed the initial pilot test results of our native speaker group. However, the statistical analysis we ran in the previous chapter included all the items ( $n = 58$ ) on which the native speaker participants were tested.<sup>55</sup> In order to compare our native speaker dataset with our L2 learner dataset, we had to rerun the analysis using a dataset that contains only those items used on the learner test. We thus removed the following items (cf. [sub-section 6.2.3](#)): IRF10 (Q10), IRF11 (Q11), IRF13 (Q13), IRF16 (Q16), IRO2 (Q18), IRO3 (Q19), IRO6 (Q22), IRO15 (Q31), OPI3 (Q35), OPI11 (Q43), OPI13(Q45) and OPI19 (Q47), and reanalyzed the accuracy and response times for native speakers.

##### *7.3.4.1. Data analysis*

It did not come as a surprise that our accuracy data was not normally distributed since we already knew that our native speakers performed at a homogenous (and very high) level. On the retained items, participants scored 100% on all questions for the IRF and OPI categories

---

<sup>55</sup> We believe Taguchi et al. (2016) based their native speaker results on the full set as well, which is why we kept that analysis as a point of comparison with Taguchi et al.

and we ran the Shapiro test (using the `stat.desc` command from the `pastecs` package in R Studio) for the IRO and Filler categories which also revealed non-normally distributed data, as seen through the values for kurtosis and skewness.

We checked the distribution of our response time dataset using the same function in R Studio and the data did not meet the assumption of normality either (the Filler category presented a positive skew and the irony category was leptokurtic). Since our data were not normally distributed, we used the same Robust one-way repeated ANOVA we used with our full datasets (using the `WRS2` package in R Studio).

Here again we used the `compute.es` package in R Studio (2021) to calculate the effect sizes and followed Plonsky and Oswald's (2014) recommendations to interpret  $d$  and  $r$ .

#### 7.3.4.2. Results

Once again, we will first present the results for comprehension accuracy of native speakers on retained items, followed by native speakers' comprehension speed on retained items.

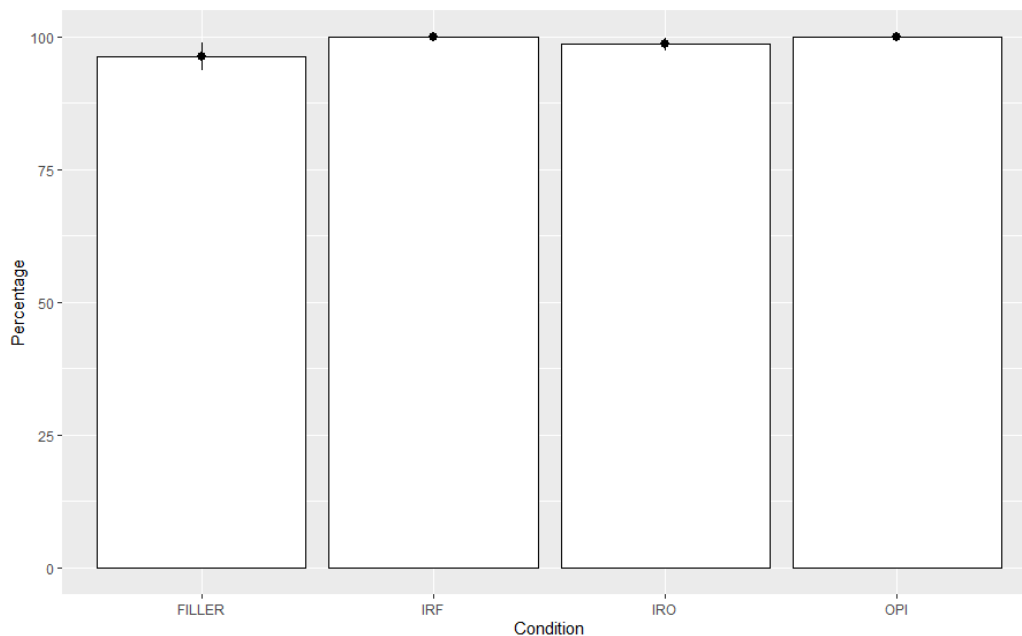
##### 7.3.4.2.1 Comprehension accuracy

After we removed the items that were not retained for the L2 test, the average accuracy scores for the native speaker group for the indirect refusals went from 98.71% to 100%, the irony category went from 97.63% to 98.56%, and the indirect opinions from 98.71 to 100%. The filler category remained unchanged since we did not remove any items from that category. [Table 18](#) shows that the native speaker participants performed at ceiling level in most item types.

**Table 18***NSs' Accuracy Scores (retained items only)*

Item types	Mean	SD	Min.	Max.
Indirect refusals	100	0	100	100
Irony	98.56	3.20	91.67	100
Indirect opinions	100	0	100	100
Fillers/Literal	96.21	6.77	80	100
Overall	98.69	2.49	92.92	100

The filler category did seem to present a wider range of scores but the mean was still very high (mean = 96.21). Using the packages `ggplot2` and `Hmisc` we created a bar chart with error bars for our accuracy scores based on participants' averages in each category. [Figure 15](#) below presents a bar graph of these scores for each category.

**Figure 15***NSs' Accuracy Scores (retained items only) - bar graph with error bars*

The statistical analysis revealed that the variable of Item type did not have a significant impact on the accuracy of responses,  $F(1, 18) = 1.4718$ ,  $p = 0.24074$ . This means that there were no significant differences between the different types of indirect/direct meaning in terms of



comprehension when looking at accuracy. The results were also consistent with the ones we obtained using the full set of data.

*7.3.4.2.2 Comprehension speed*

While accuracy score was the main factor on which we based selection of items to be removed from the L2 test, response times were also taken into account and some of the removed items also had higher times as compared to the retained items. As a result, the response times (on correct answers) for the three indirect meaning categories went down and the average response time across all four item types went from 7.29 sec to 6.96 sec.

The average for indirect refusals went from 7.58 sec. to 7.06 sec., the irony went from 7.52 sec. to 7.15 sec. and the indirect opinions from 6.89 sec. to 6.47 sec. Since we did not remove any items from the filler category, the data remained unchanged for that condition. [Table 19](#) presents the descriptive statistics for the native speaker group comprehension speed on the retained items (items that were selected for the L2 test).

**Table 19**

*NSs’ Response Times on correct answers (retained items only)*

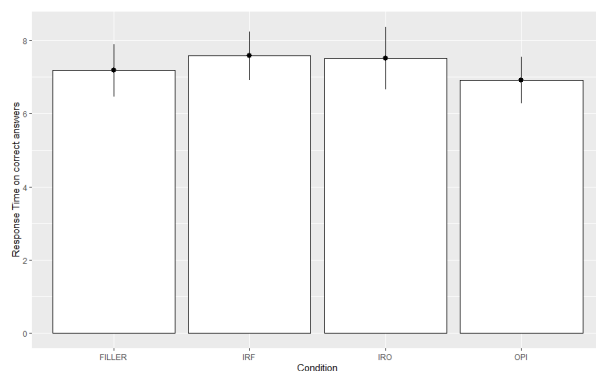
	Mean	SD	Min.	Max.
Indirect refusals	7.06	1.55	4.06	12.29
Irony	7.13	2.18	3.67	12.20
Indirect opinions	6.47	1.54	3.69	10.77
Fillers/Literal	7.18	1.88	3.86	12.87
Overall	6.96	1.79	3.82	12.03

This difference in means between the full dataset (cf. [Table 12](#)) and [Table 19](#) (above) seemed fairly small (0.33 sec.) so we did not anticipate that it would change the results of our statistical analysis. However, it is interesting to compare the bar charts that were created using the full dataset ([Figure 12](#)) vs. the dataset with the retained items only ([Figure 16](#)).

Below is a side-by-side comparison of both charts.

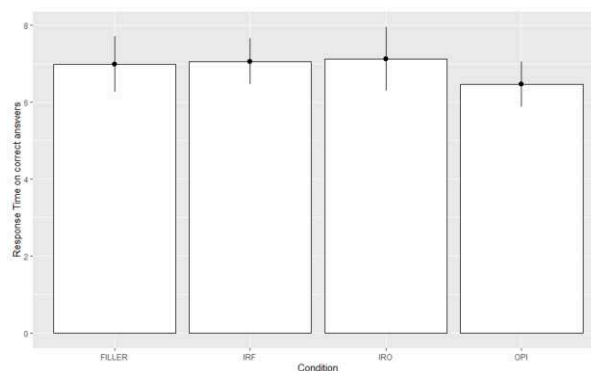
**Figure 12 NSs Response Times**

*NSs' Response Times on correct answers – bar graph*



**Figure 16 NSs Response Times on correct answers (retained items only)**

*NSs' Response Times on correct answers (retained items only) – bar graph*



We can see here that in [Figure 12](#), the response times for the filler category were on average lower (i.e., faster) than that for the indirect refusals and the irony, and the indirect opinion category had the lowest response time. Our initial statistical analysis (cf. [sub-section 6.3.2.2](#)) had revealed a statistical difference in response time, but the post hoc test only showed a difference of negligible effect size between the OPI and IRF conditions. [Figure 16](#) (with retained items only) still shows the OPI category as having the lowest (i.e., fastest) response time average, but the filler category is now much closer to the IRF and IRO category. Irony here seems to have required the most amount of time for participants to understand.

We performed a Robust one-way repeated ANOVA on our dataset, which revealed no significant difference in native speakers' comprehension speed across all four categories,  $F(2.19, 39.42) = 2.6185$ ,  $p = 0.08092$ . This confirms that our native speaker group performed in a consistent manner regardless of the item type on those items that were retained for the L2 learner test.

#### *7.3.4.2.3 Comparison of native speakers and L2 learners*

Expectedly, our L2 French learners' comprehension of indirect meaning was not as good (72.64%) as that of our native speakers. Indeed, our native speakers reached 99.52% accuracy on the same indirect meaning items.

We found no significant difference in accuracy and speed in our new native speaker dataset, which means our native speakers did not have any more issues with indirect meaning than they did with literal meaning. This was different from what Taguchi et al. found with their native speakers. The fact that our native speakers had no issues understanding indirect meaning, but our L2 learners struggled with certain types of indirect meaning (particularly irony), indicates that comprehension difficulties observed in our study were likely related to L2 learning/processing and not simply an issue with indirect meaning being more challenging (for native speakers and L2 learners) as Taguchi et al.'s findings seem to suggest.

## Part II - Conclusion

As part of our replication study, we wanted to investigate L2 French learners' comprehension of indirect meaning and find out whether Taguchi et al.'s (2016) unexpected finding with regard to conventionality could be confirmed. Our analysis of our quantitative data allowed us to increase the confirmatory power of Taguchi et al.'s study in the sense that our L2 French learners showed no significant differences in the comprehension of indirect refusals (which were categorized as conventional by Taguchi et al.) and indirect opinions (which were categorized as nonconventional). Our findings with regard to irony comprehension being more challenging for L2 learners also corroborated Taguchi et al.'s findings as well as research by previous authors.

The inclusion of the filler items (literal meaning items) in our analysis allowed us to show that not all types of indirect meaning are necessarily more challenging to understand. Indeed, indirect opinions and indirect refusals were not actually more difficult to comprehend for our L2 French learners than literal items. This was something Taguchi et al. did not investigate in their study<sup>56</sup> (they did not include the literal items in their analysis), which we thought was lacking. Indeed, it would be hard to explore the extent to which L2 learners understand indirect meaning without comparing with direct/literal meaning comprehension.

In various sections of [Part I](#) and [II](#) of this dissertation, we have also questioned Taguchi et al.'s categorization of indirect meaning types as conventional or non-conventional. Looking into the processes by which our L2 learners infer intended meaning may show whether the presence of certain conventional patterns did indeed help them reach the correct interpretation of the indirect meaning utterances, as well as provide some explanations as to why irony was significantly more challenging or why indirect opinions did not seem to pose too many issues. [Part III](#) of this dissertation will focus on the retrospective interviews we conducted for this very purpose.

---

<sup>56</sup> We should note however that Taguchi et al.'s accuracy results on indirect refusals and opinions were even higher than ours (around 90%), suggesting that the indirect nature of speech acts also caused little comprehension difficulty for their learners.

# **PART III**

## **RETROSPECTIVE**

### **VERBAL**

#### **INTERVIEWS**

## Introduction

Part III of our dissertation will concentrate on the qualitative part of our research which attempts to answer our third research question:

### **3) What inferential strategies do L2 French learners use and how successful are they?**

Following Taguchi et al.'s study, we conducted retrospective verbal interviews among two sub-groups of participants. The purpose of the follow-up interviews was to gain some understanding into the comprehension processes used by these participants, identify elements of difficulty, and contrast the strategies used and cues attended to by low-scoring and high-scoring participants for the different categories of indirect meaning.

[Chapter 8](#) will present the methods related to the interviews: interview participants ([section 8.1](#)), test item selection ([section 8.2](#)) and interview procedures ([section 8.3](#)), following with a section on our coding protocol ([section 8.4](#)) and data analysis ([section 8.5](#)). In [Chapter 9](#), we will focus on the interview results and discuss our findings. We will first look at possible differences between low-scoring and high-scoring groups ([section 9.1](#)), trying to figure out what caused each group difficulties: was one group more successful in their interpretation and use of strategies? When the sentence-level meaning of the target utterance was not accessible, were strategies used to circumvent L2-specific deficiencies (e.g., inadequate lexical knowledge, poor listening abilities, etc.)? We will then analyze the impact of the different types of indirect meaning on strategy selection ([section 9.2](#)), and try to investigate the possible impact of conventionality. Were differences in accuracy scores and comprehension speed on the in-class test explained by the strategy selection patterns identified in the verbal reports?

## Chapter 8 - Methods

While our methodology tried to follow Taguchi et al.'s (2016) protocol, there were instances when we did not exactly know the specific procedures they used, due to the fact that they did not explicitly report on them in their paper. For example, we did not have access to their transcript, the full list of test items they used with the interview participants and the data they collected. Additionally, very little information was shared regarding coding in Taguchi et al., but we will present our coding protocol in [section 8.4](#).

### 8.1. Participants

Following Taguchi et al.'s (2016) protocol, our participants were divided into three groups based on their test accuracy scores: one third (11 participants) represented the highest-scoring participants, one third the lowest-scoring participants (11 participants) and the last third was the remaining 11 participants whose accuracy scores fell in the middle (between top- and bottom-scoring groups). Taguchi et al. collected qualitative data from four participants from the top-scoring group and four from the low-scoring group (selected at random) during follow-up retrospective verbal interviews, and we aimed to follow the same system.

Two participants in our lowest-scoring group had indicated that they did not want to be contacted for a follow-up interview on the online survey. We then decided to reach out to the remaining four participants having received the lowest scores, to schedule the interviews that were to happen about three weeks after the first round of data collection. One participant did not reply to our request, so we selected two additional participants and contacted them for a follow-up interview. We decided to include an extra participant in order to guarantee having at least four respondents. In the end, we scheduled follow-up interviews with five low-scoring participants.

A similar process was used for the high-scoring group. All 11 participants had volunteered for the follow-up interviews and we contacted five of the top-scoring students (we decided to exclude the student who was also a Spanish speaker and ranked #5 on the accuracy score). One participant did not reply. Four follow-up interviews were scheduled during the third week after the data-collection activity. In total, nine participants completed the retrospective verbal

interview: four males and five females, 28% of our sample). [Table 20](#) shows all participants' accuracy scores (out of 46 points and in percentage), ranked from lowest to highest. The participants who participated in the follow-up interviews are highlighted in grey.

**Table 20**

*Participants' accuracy scores*

<i>Low-scoring</i>						<i>High-scoring</i>		
Participant	Accuracy score		Participant	Accuracy score		Participant	Accuracy score	
1	24	52.17%	10	32	69.57%	23	38	82.61%
2	26	56.52%	15	32	69.57%	24	38	82.61%
4	28	60.87%	13	33	71.74%	25	38	82.61%
5	28	60.87%	14	33	71.74%	26	39	84.78%
7	28	60.87%	17	35	76.09%	27	40	86.96%
3	29	63.04%	18	35	76.09%	28	41	89.13%
6	29	63.04%	16	36	78.26%	29	42	91.30%
8	30	65.22%	20	36	78.26%	30	42	91.30%
11	31	67.39%	21	36	78.26%	31	42	91.30%
12	31	67.39%	19	37	80.43%	32	44	95.65%
9	32	69.57%	22	37	80.43%	33	45	97.83%

The overall test accuracy rate (on the in-class test) for the low-scoring participants selected for the interviews was 58.69% and the average for the high-scoring participants we interviewed was 92.39%.

It is important to note at this point that the original plan was to conduct the interviews on campus. However, due to the Covid-19 health crisis, the campus was shut down and we were unable to hold in-person interviews. The details of the interview administration will be discussed in the Procedure section below ([section 8.3](#)).

## 8.2. Item selection

In their study, Taguchi et al. focused on nine experimental items (three from each category) that “functioned well to discriminate between the high and low-scoring groups” (p. 687). In order



to replicate their method, we averaged the scores for each of the items for the 11 members of the low-scoring group and repeated this process for the 11 people that made up the high-scoring group, in order to identify the questions that presented the highest difference. In addition, we decided to include the literal meaning items (distractors in Taguchi et al.'s study) as a fourth category in the retrospective interviews.<sup>57</sup> Therefore, where Taguchi et al. reviewed nine items, we reviewed 12 (three from each of the following categories: indirect refusals, irony, indirect opinions and literal meaning/fillers).

IRF 2, 6 and 9 (Q2, 6, 9) were selected for the indirect refusals based on the highest difference between high-scoring and low-scoring group. For example, the average score on IRF 2 was 0.182 for the low-scoring group and 1.00 for the high-scoring group, which means the difference was 0.818. The difference for IRF 6 was 0.636 and for IRF 9 it was 0.545. IRF 14 also had a difference of 0.545 but we decided to use IRF 9 over IRF 14 because the difference for IRF 9 was greater when we only looked more specifically at the scores of the participants that were selected for the follow-up interviews. The same principle was used for other items which had an identical score. IRO 5, 8 and 14 (Q21, 24 and 30) were selected for the irony category, OPI 2, 7 and 8 (Q34, 39 and 40) were chosen from the indirect opinions and Fill 4, 6 and 7 (Q52, 54 and 55) were used for the literal meanings. The complete table of items with low/high-scoring mean and difference between them can be found in [Appendix O](#) (the items that were selected will appear in yellow). The item selection was done based on the answers provided by the entire low and high-scoring groups (as defined by Taguchi et al. as the “bottom third” and the “top third,” p. 683), not just the few participants who were selected for the retrospective verbal interviews, which, of course, led to some minor discrepancies in accuracy scores in the groups of participants that we interviewed as compared to the overall groups used for the item selection. However, when we looked at how the participants that we selected for the interview performed, the difference between low and high-scoring on the selected items was very clear and consistent with the two groups: test scores on the interview items was an 85.42% for the high-scoring interview participants, and the low-scoring interview group only reached 33.33% accuracy.

---

<sup>57</sup> The data from the literal items were ultimately not analyzed because we were focusing on identifying inferencing strategies used to comprehend indirect meaning, which would not apply to literal (direct) meaning comprehension.

### 8.3. Procedure

As mentioned in [section 8.1](#), we had initially planned on holding these interviews on campus. Since our campus closed due to Covid-19 and we were not able to meet with participants in person, we decided to conduct the interviews via the video-conferencing platform Zoom. Participants had only consented to audio recording as part of the initial IRB process; therefore, volunteers were notified that they did not need to turn on their video if they did not feel comfortable being video recorded as the audio would be enough. We followed Taguchi et al.'s interview protocol. Each item was presented one by one. We played the video using the screen-sharing function and then displayed the multiple-choice statements on the screen so the participant could select their answers. We then asked the same type of questions for each item: "Why did you choose this answer?," "What were you thinking when you chose this answer?." It is important to note here that Taguchi et al. did not provide a lot of details regarding their interview protocol, and without having access to the complete interview transcripts, we do not know whether her team may have asked more specific questions than those provided as examples in their paper, or whether they followed up with additional requests for details.

In some instances, after recording the initial comment from the participants, we asked them to elaborate on some part of their answers to get additional feedback. For example, participants may have been explicitly asked to discuss specific cues they may have attended to ("Was there anything specific you may have noticed in the video in the way the person was acting? Any facial expressions or gestures?"). Some participants seemed to need more probing and initially did not seem to understand what we meant when we asked about their comprehension strategies ("How did you get to that answer?"). While we realize that the additional questions may have also influenced the participants' feedback later in the interview, it was important for us to try to understand their thought process better in order to answer our research questions. To that effect, on several occasions, we provided additional explanations to the students who were not able to understand the video (translation of the dialogue) to see if they could then interpret the intended message properly once they had access to the sentence-level (linguistic) meaning. We will discuss the potential impact of these interview questions and strategies in [Chapter 9](#).

Our overall interview lengths ranged from about 20 to 30 minutes each and we collected 12 reports for each of the nine participants, for a total of 108 items. Our interview sessions were longer than Taguchi et al., likely due to the fact that we reviewed additional items (the

fillers/literal meanings questions), and because the beginning and end of the interviews also included some talk that was not relevant to our study (in particular with regard to the pandemic situation, transition to online classes, etc.).

Each interview was then manually transcribed by the researcher (using the Zoom auto-generated transcript as a base).

#### 8.4. Coding choices for the replication study

For the purpose of our replication, we decided not to include one of the lowest-scoring participants in our analysis, since we interviewed a total of five for that group, and Taguchi et al. collected data from four participants per group. Given that Participant 1 experienced some technology-related issues during the interview (they used a mobile device and their poor connection led to them having trouble hearing and seeing the videos, which required us to play the videos multiple times), we excluded their answers from our coding. Data from the following participants were retained for our analyses (refer to [Table 20](#) for details on these participants' overall accuracy):

- Low-scoring: Participant 2, 3, 5, and 7
- High-scoring: Participant 28, 29, 31 and 33

Taguchi et al.'s main goal in doing their retrospective verbal interviews was to gain some understanding into the learners' inferencing mechanisms, "whether L2 listeners' attention is bound to linguistic input, or whether they attend to nonlinguistic visual cues such as facial expressions and gestures during comprehension" (p. 682). Prior to their 2016 paper, Taguchi had already investigated strategies that learners use to access indirect meaning (Taguchi, 2002, 2008a). In the 2016 paper, Taguchi and colleagues also added the multimodal aspect to their study: Taguchi's previous studies only used auditory input, but the multimedia test allowed them to include visual cues as possible strategies to assist comprehension. Drawing from previous research (in particular Attardo 2000b; Attardo & Poggi, 2003; Attardo et al., 2013), they explained that "comprehension is not merely the decoding of linguistic input; rather, it is the global process that involves the use of all available cues, both linguistic and nonlinguistic (visual and auditory cues) to arrive at the meaning" (p. 682). While they presented some numerical data (e.g., they list the number of strategies that were used—classified according to

the type of strategies and the participant group—whether high or low-scoring), they also aimed to investigate L2 indirectness comprehension through “qualitative analyses of participants’ inferencing processes,” in order to discover “what causes comprehension difficulty” (p. 683). We conducted our own interviews with this objective in mind.

Taguchi et al. identified five categories of strategies, which they based on previous research (Taguchi, 2002) on the inference process for comprehending indirect meaning: (Taguchi et al. 2016, Table 4-p. 690):

### Figure 17

*Categories of Inferencing Strategies and Sample Interview Excerpts from Taguchi et al. (2016)*

**Auditory cues:** Identifying vocal qualities (intonation, stress, pause, tone, or speech rate) to infer the speaker’s intention.

*“I thought that her tone sounded a little bit more negative than positive. So that’s why I chose number 3 and not number 4.”*

**Visual cues:** Identifying visual cues (e.g., facial expressions and gesture) to infer the speaker’s intention.

*“u:h he made u:h a face at her haircut that look like dislike or disgust so I was like, nah he doesn’t like it.”*

**Background knowledge:** Referring to life experience and global knowledge to make inferences.

*“If your back hurts more after than before then it’s not a good thing and you probably don’t like it.”*

**Key word inferencing:** Picking up a key word and associating it with meaning.

*“I’m pretty sure I heard *sufre* in the last sentence [laughs] and (2.0) I think she meant like (2.0) suffering that the trip was like a bad thing.”*

**Logical reasoning:** Working deductively toward the intended meaning.

*“Marcela said that she didn’t have enough money to even buy a jacket. So she can’t go. Uhm. traveling with him to London.”*

As mentioned previously ([sub-section 4.3.3](#)), some of these categories were problematic. Indeed, it seemed to us that the dialogue sample provided by Taguchi et al. for the “background knowledge” category could also be an example of logical reasoning (with deductive pattern “if... then”). Similarly, a certain degree of background knowledge appears to be necessary to interpret the example used for the “logical reasoning.” One would need to know that traveling

to London is expensive (and comparatively more expensive than buying a jacket) to reach the correct interpretation. It is however possible that whenever there was an overlap, Taguchi et al. coded both strategies for these examples.

Looking at our transcriptions, we coded each strategy as they appeared in the initial answer provided by our participants for each verbal report. We did not code the verbal reports focusing on literal items since our analysis (and that of Taguchi et al.) focused on inferencing strategies for accessing indirect meaning and therefore would not be applicable to literal meaning. Answers or explanations that were given after we provided some additional assistance (e.g., translation of part of the dialogue) were not included. In our spreadsheet we made note of whether the strategies ultimately resulted in the selection of the correct answer or not, and also included the result from the in-class test (whether participants were successful in selecting the correct MCQ option during the first part of the study).

As we were coding the interview transcripts, it became obvious that some of the categories used by Taguchi et al. were either overlapping, needed to be redefined, or seemed to not cover some other strategies we saw in our data. For that reason, we chose to include a sixth category we labeled “Other” which we will also discuss in [sub-section 8.4.5](#). Given that we used the same bottom-up approach Taguchi et al. used in their study, it is not surprising that we may have similar categories, but because our coding process (like Taguchi et al.) was data-driven, it allowed for the possible identification of additional strategies. Another thing we should keep in mind is the fact that we were only able to code inferencing strategies that were verbalized by participants, and it may very well be the case that participants did not verbalize all processes.

[Figure 18](#) presents the definition of the different strategies we found in our interviews along with sample excerpts from our interviews. These include the five strategies identified by Taguchi et al. and four additional strategies that were uncovered in coding our data. Some modifications were made to the original strategies proposed by Taguchi et al. In the section that follows, we will detail each strategy, explain our operationalization, and discuss difficulties we encountered. In the following figure, each strategy is color-coded: the same colors will also be used in the excerpt transcripts we will provide throughout Chapter 8 and 9 for ease of identification.

## Figure 18

Categories of Inferencing Strategies with example from our interviews

**Auditory cues:** inferring the speaker's intended meaning based on vocal qualities (intonation, stress, pause, tone, or speech rate).

*"their voices were kind of flat."* [IRF6 – Participant 5 – low-scoring]

*"she spoke in a sarcastic tone."* [IRO14 – Participant 33 – high-scoring]

**Visual cues:** inferring the speaker's intended meaning based on visual non-verbal cues or body language (facial expressions, gestures, etc.)

*"I just definitely saw her, like, eye roll thing at the end."* [IRO5 – Participant 2 – low-scoring]

*"[...] honestly what got me to that particular answer was the gesture at the end."* [IRF6 – Participant 5 – low-scoring]

**Background knowledge:** inferring the speaker's intended meaning based on one's life experience or global knowledge (including knowledge of the L2 cultural conventions).

*"Friends will be like, you know, "must be nice," kind of thing to each other. And even if, like they're happy for each other. They'll be like, oh yeah, whatever."* [IRO5 – Participant 2 – low-scoring]

*"[...] the fact that he said that he's going to all these different cities like that's not that's that's a pretty good deal."* [IRO5 – Participant 31 – high-scoring]

**Keyword:** inferring the speaker's intended meaning by relying on the presence or absence of a specific keyword (or set of keywords) or lexical items in the dialogue.

*"first you could hear them say hot chocolate at some point. I remember them saying "chocolat" and like "chaud". [...] then the last word he says is "maison", which is "house"."* [IRF 2 – Participant 7 – low-scoring]

*"Um, well, I heard "Tuesday." And then, um, I heard "n'ai," so I figured it was a negative statement"* [IRF6 – Participant 7 – low-scoring]

**Logical reasoning:** recognizing literal meaning and working deductively toward the intended meaning.

*"They're not going to meet on Tuesday, because he says he doesn't come to the University on Tuesday."* [IRF6 – Participant 33 – high-scoring]

*"So, because obviously, you know, if your answer is that she thinks that it's cold, but then she says it's a sauna, that doesn't make sense."* [IRO14 – Participant 3 – low-scoring]

### Other:

**Setting/situation:** inferring the speaker's intended meaning based on situational/contextual elements present in the video or description of the scene.

*"Because the context of them being in the cafeteria."* [IRF2 – Participant 5 – low-scoring]

**Familiarity with the characters:** expecting the characters to behave or to express themselves in a certain way (sarcastically, indirectly, etc.) based on familiarity with their attitude or behavior from other videos.

*“I noticed that a lot of times, she would kind of give him a hard time in a lot of videos so it kind of made me like suspicious that like she might be sarcastic.”*

[IRO5 – Participant 31 – high-scoring]

*“because if she would have liked it, like in the previous one, she would have been like smiling or something”* [OPI7 – Participant 29 – high-scoring]

**Adjacency pair rule:** Using knowledge of conversation structure (e.g., adjacency-pair rule) and how certain turns in conversation create expectations for the next turn.

*“He asked like “how was the massage?” after she got it, and it wasn't positive, so it seemed like her back could have hurt.”* [OPI7 – Participant 29 – high-scoring]

*“Maybe because of her response to him and saying like, no, like, I'm, I'm good, because he was kind of like questioning her.”* [IRF 2 – Participant 2 – low-scoring]

**Process of elimination:** specifically referring to the multiple-choice questionnaire options and proceeding by elimination to infer the intended meaning.

*“We're taught the process of elimination in elementary school like, you'll see a couple of answers that you can immediately rule out and some answers can have context to others.”* [IRO8 – Participant 5 – low-scoring]

*“I know like he wasn't available on Tuesdays. So (3.0) it's not three. It's not four. Uh, I would choose Number one.”* [IRF6 – Participant 2 – low-scoring]

In some instances, participants combined different strategies centering on a specific element of the video. For example, they may have noticed a certain negative facial expression that led them to understand the ironic intent of the utterance, and also expressed their logical reasoning to explain the incongruity between what was said and what the person was showing with non-verbal cue. We coded for all these instances of strategies (which means that one single answer could be coded for multiple strategies<sup>58</sup>). Examples of the use of combined strategies will be provided to illustrate the categories in the next five sub-sections (8.4.1 through 8.4.5) which present all the above strategies and discuss the coding decisions we made and challenges we encountered.

#### *8.4.1. Visual and auditory cue strategies*

The first two strategies in our table are very close to those of Taguchi et al.'s. Auditory cues were coded as such when participants mentioning hearing a specific intonation or tone, or described the way the characters “sounded” in the video. Visual cues were coded when

---

<sup>58</sup> This follows Taguchi et al.'s coding process: Based on the total number of strategies they identified (100 for 72 verbal reports), we can infer that they did find instances when more than one strategy was used for a specific answer.

participants mentioned noticing specific non-verbal cues (gestures, facial expressions, body language) that led to the intended meaning. While these first two strategies (visual and auditory) were fairly straight forward to code, we sometimes had to ask for clarification when participants used ambiguous language, as illustrated in [Excerpt 1](#).

### Excerpt 1

*Video Q24 – IRO8 - Participant 29 [high-scoring]*

**Participant 29:** She's kind of like smiling at him, she's kind of like snarky at the end.

**Researcher:** Yeah

**Participant 29:** Like she's shaking her head.

**Researcher:** Yeah. Yeah, so she's shaking her head. And when you say she's being like snarky. Is it just, are you just basing that on what she's saying, or is there something else...

**Participant 29:** it's because of the way she's looking at him. Her eyes are like squinted

**Researcher:** Mhhmm

**Participant 29:** Like they're like pointed directly at him, but they're not like a genuine smile, you know.

In [Excerpt 1](#), Participant 29 clearly noticed some visual cues but we were not sure what they meant when they said the character was being “snarky” as this could have been based on the dialogue, or a different intonation Participant 29 could have picked up on. Her reply to our request for clarification indicated that she only attended to visual cues in this specific excerpt.

While participants usually noticed visual cues from the last person who spoke in the video (since the last comment was the focus of the study), they sometimes mentioned non-verbal cues from the other character in the video, denoting the interactive aspect of the conversation. For example, Participant 29 reported relying on the way the female character looked to infer the indirect refusal: “Um, he said that he had like prior engagements, pretty much. And **then she looked kind of disappointed**” (Q6, IRF6, Participant 29). We coded all instances of non-verbal body language that originated from either character in our video as instances of visual cue strategies (as long as they were directly linked to the target utterance).

Participants did not provide detailed descriptions as to the auditory cues they attended to. Usually they referenced a specific “tone” or “intonation” as shown in [Excerpt 2](#).



## Excerpt 2

*Video Q30 – IRO14 - Participant 33 [high-scoring]*

**Researcher:** [showing the MCQ options on the screen] Alright. What are you thinking?

**Participant 33:** She still thinks it's cold in class and you can tell because she raised her eyebrows and spoke in a sarcastic tone and said, “Yeah, it feels like we're in a sauna.”

Note that the above excerpt was also coded for two other strategies (visual cues and logical reasoning—the latter will be discussed in [sub-section 8.4.4](#))

### *8.4.2. Background knowledge strategy*

Background knowledge was coded whenever participants verbalized relying on personal life experiences or common knowledge as previously stated, but also included instances of participants' meta-analysis of certain communicative strategies, which were often derived from the participants' exposure to similar situations or knowledge of L2 cultural conventions. In our view, it includes macro- and micro-social norms, commonsense assumptions and factual information that are part of the contextual source that Yus (2000) labelled “encyclopaedic information.” For example, [Excerpt 3](#) shows that Participant 2 was able to correctly interpret an irony item by drawing a parallel with English and based on her personal knowledge of people's interactions:

## Excerpt 3

*Video Q21 - IRO 5 - Participant 2 [low-scoring]*

**Researcher:** Just remember that, here we're focusing on her last comment, what she said at the very end. Um: So I don't know if you if you understood the situation. What, what, what did you get from that situation? They're talking about his trip:

**Participant 2:** Um, I just definitely saw her, like, eye roll thing at the end and then she's just kind of like, um: like, uh, just kind of not caring about it. U:h

**Researcher:** Yeah. So that might help you with the answer, right?

So number one: Alicia thinks that Richard life is pretty good, or she thinks that Richard is quite unhappy [pointing option #2 on the screen]. Alicia and Richard plan to make the trip to Europe together [pointing option #3 on the screen], or she thinks Richard went through really hard things in his life [pointing option #4 on the screen].

**Participant 2:** I think number one because like friends will be like, you know, “must be nice,” kind of thing to each other. And even if, like they're happy for each other. They'll be like, oh yeah, whatever. You know, you [giggles]

**Researcher:** Yeah. Yeah, absolutely

The last comment in [Excerpt 3](#) was coded as an example of background knowledge, but we also coded for visual cues as Participant 2 mentioned seeing the female character roll her eyes. Even though Participant 3 later admitted that she did not understand the literal meaning of the target utterance, she was able to draw from her personal experience with interpersonal interactions between friends to figure out the ironic intent of the utterance.

Another example of reliance on background knowledge that we observed in our data, and that Taguchi et al. also discussed in their paper (2016, p. 692-693) came from the same dialogue [Excerpt 3](#) focuses on: the male character, Richard, discusses his upcoming vacation and lists all the cities in Europe that he is planning to visit, to which the female character, Alicia replies “*Oh non, arrête, tu as la vie dure!*.” As Taguchi et al. state (p. 693—note that the character named Ricardo in Taguchi et al. was named Richard in our dialogue):

To detect this underlying meaning, learners needed to use their background knowledge about the cities Ricardo was visiting. Once they understood that these cities are all ideal places to visit, they recognized the mismatch between Alicia’s proposition (Ricardo was suffering) and intended meaning (Ricardo was looking forward to his trip).

Although all the participants who reached the correct interpretation for this particular dialogue and understood the sarcastic comment may have actually relied on their background knowledge to be able to notice the incongruity between what Alicia said and the situation (Richard visiting all these European cities), we could only code for this strategy when the participants actually verbalized this process (e.g., Q21, IRO 5, Participant 31: “**the fact that he said that he's going to all these different cities like that's not that's that's a pretty good deal.**”).

#### 8.4.3. *Keyword strategies*

Taguchi et al. (2016) coded for keyword inferences when participants picked up on a specific keyword and associated it with a certain meaning, which could either correspond to the intended meaning or not. This type of strategies appeared to be typical when a person did not completely understand a dialogue or sentence, and they tried to elaborate meaning on the basis of the few words they understood (or thought they did). We noticed that in our interviews, participants also picked up on words present in the dialogue, but at times, they mentioned relying on key words that were not actually in the dialogue, but they thought they heard in the video. We also

coded these instances as keyword inferencing. Additionally, participants would sometimes base their answers on keywords that they specifically did NOT hear in the dialogue. That is to say, as illustrated in [Excerpt 4](#), participants explained tuning in to a particular word that may have clued them into the indirect meaning, and basing their answer on the fact that the word was not, according to them, in the dialogue.

As a result, we defined our category of keyword strategy as “inferring the speaker’s intended meaning by relying on the presence or absence of a specific keyword (or set of keywords) or lexical items in the dialogue.” [Excerpt 4](#) is an example of a keyword strategy that was coded based on the absence of a keyword. It should be noted that this sub-set of keyword inferencing strategies was usually combined with the “elimination process” strategy (which we will discuss later). We felt it was important to code for both, since the elimination process strategy was sometimes combined with other strategies (visual cues/auditory cues, etc.), and since the participants were still directing their attention to specific keywords, it seemed appropriate to include these instances in the keyword-relying strategy. [Excerpt 4](#) illustrates this particular strategy: while the participant was able to understand the gist of the beginning of the dialogue, he could not understand the target sentence. However, he turned his attention to specific keywords that he remembered not hearing in the video (therefore eliminating certain answers and interpretations from the multiple-choice questionnaire).

#### **Excerpt 4**

*Video Q30 – IRO14 - Participant 7 [low-scoring]*

**Researcher:** Alright, so what are you thinking for that one?

**Participant 7:** I would say four.

**Researcher:** Okay.

**Researcher:** So how did you get to that?

**Participant 7:** Um, to be honest, I don't remember what I put but I think I would have put either one or four because...

**Researcher:** Yeah

**Participant 7:** they're saying that it's cold in class. And she's like, “Yeah, that's why I bring my jacket” and things like that. Um, But at the end, I can't really make out exactly what she says whether she, I don't think she feels... Yeah, I actually want to change it to one. [giggles]

**Researcher:** Okay, that's good. And so again, based on, based on just the beginning of the conversation or...

**Participant 7:** Yeah, just I switched it based off of, like, [I never hear like either of them say really warm or anything like that.](#)

#### 8.4.4. Logical reasoning strategies

While we believe that most inferencing processes require some amount of logical reasoning (cf. [section 3.3](#)), within the framework of our close replication of Taguchi et al. (2016), we wanted to make sure we coded for strategies of logical reasoning in a way that was similar to what Taguchi et al. did. Therefore, we tried to limit ourselves to those instances when our participants verbalized using logical reasoning that stemmed from literal meaning as this was what we understood Taguchi et al.'s logical reasoning strategy to be based on (cf. sub-sections [4.3.3](#) and [4.4.2](#)). [Excerpt 5](#) shows a clear example of logical reasoning with a cause and consequence pattern.

#### Excerpt 5

*Video Q9 - IRF 9 - Participant 28 [high-scoring]*

**Researcher:** OK, so here are the options [showing the MCQ on screen], what do you think?

**Participant 28:** That she will not try tomorrow with Eric at 7pm.

**Researcher:** Okay. Perfect. So how did you get to that answer?

**Participant 28:** Again, the dialogue, just **the fact that she has a concert to get to at 7:15 so obviously she wouldn't be able to make a dinner at seven o'clock.**

**Researcher:** Very good. Any other things that you attended to

**Participant 28:** No just the dialogue that's it.

**Researcher:** Okay. Thanks.

Just like the previous categories of inferencing strategies, there were instances when our participants combined the logical reasoning strategy with other sources, such as visual, auditory, or background knowledge, as will be shown in our posthoc analysis in [Chapter 9](#).

#### 8.4.5. Additional Strategies

During our coding phase, we came across what we believe to be four additional strategies that were not listed by Taguchi et al. One of the reasons why we thought it was appropriate to expand their strategy list for our replication was because, while they indicated their coding was based on Taguchi (2002), there were some differences between the strategies they presented in their 2016 paper and the ones that appeared in Taguchi (2002) (see sections 4.3 and 4.4). Additionally, the coding process they described was data-driven, which means the identification of the strategies did not start from a pre-established list of strategies but rather these strategies

emerged during the coding phase. Since we followed a similar approach, we knew we may encounter strategies that were not manifest in Taguchi et al.'s (2016) data.

We will now take a closer look at these other strategies we identified in our verbal reports, which are setting/situation, familiarity with the characters, adjacency pair rule and process of elimination.

#### 8.4.5.1. Setting/situation cues

In [sub-section 8.4.1](#), we mentioned that the definition of “visual cue” strategies provided by Taguchi et al. mainly focused on non-verbal body language that could be seen in the video, such as facial expressions and gestures, in particular when associated with the target sentence of the dialogues (where the indirect or direct meaning sentence on which we were focusing were located). Therefore, it is important to note that per ours and Taguchi et al.'s description, the “visual cue” strategy category did not include ALL type of visual elements that could be seen in the videos. This distinction was one of the reasons we decided to create this new category of contextual or situational strategies based on the setting or situation. We defined this category as follows: “inferring the speaker’s intended meaning based on situational/contextual elements present in the video or description of the scene” (cf. [Fig. 18](#)). Indeed, our participants sometimes mentioned relying on elements they had seen in the videos to reach their interpretation but these elements were not examples of non-verbal body language. Examples of this were provided in [Fig. 18](#), and overall, these types of contextual cues were usually associated with the physical setting of the scene (characters being in a cafeteria, or wearing specific clothes) as well as details provided in the video description. This type of information is what Yus (2000, p. 7) would refer to as the “mutually manifest physical environment” (cf. [sub-section 2.6.4](#)).

#### 8.4.5.2. Familiarity with the characters

This strategy only appeared twice in our data. We defined it as “expecting the characters to behave or to express themselves in the certain way (sarcastically, indirectly, etc.) based on familiarity with their attitude or behavior from other videos” (cf. [Fig. 18](#)). We believe it is similar to what Yus (2000) referred to with the “addressee’s background knowledge of addresser’s biographical data” (p. 8-9). In our verbal reports, there were a few times when

participants mentioned noticing a pattern of behavior from characters who were particularly sarcastic. Of course, this is likely due to the fact that a disproportionate number of our items focused on indirect meaning (as opposed to literal meaning) and that the female actress in our videos did end up playing the role of a character who often used sarcasm. Each video was supposed to present two different characters but, following Taguchi et al., the same two actors in our study played all the characters in the videos. We may argue that this was a design flaw in our study which influenced the participants' inferencing process. Yet, we know that even in real life situations, our familiarity with situations or people set certain expectations for the way we will process and comprehend utterances. For example, Yus (1998) describing this category of contextual source, state that "if we know that someone is prone to using irony in any conversational interaction in which he or she is engaged, this information will make us be on our guard against possible ironic messages underlying this person's utterances" (p. 5). In [Figure 18](#), we provided two examples of such strategies used by participants.

#### 8.4.5.3. Adjacency pair rule

We have discussed the mechanisms of the adjacency pair rule in Chapter 2 ([section 2.7](#)) and also mentioned that while Taguchi et al. did not observe this strategy in their data, Taguchi (2002) did note that some of her participants were referencing this pattern of conversation in their inferencing process. We described this strategy as "Using knowledge of conversation structure (e.g., adjacency-pair rule) and how certain turns in conversation create expectations for the next turn," after Taguchi (2002). We provided some examples in [Figure 18](#): only instances in which the participants referred to the previous utterance in identifying a communicative pattern were coded for this strategy. For example, in her verbal report focusing on OPI7 (in which the male character asked the female character how her massage was and she indirectly answered that it was not good by saying "*Je n'avais pas mal au dos avant ce massage!*"), Participant 29 (high-scoring) did not fully hear the target utterance but observed the adjacency pair in the request for opinion: "He asked like 'how was the massage?' after she got it, and it wasn't positive, so it seemed like her back could have hurt" (Q39, OPI7, Participant 29 – High-scoring). The question from the male character primed the hearer to expect some kind of evaluative judgement on the massage.

#### 8.4.5.4. Process of elimination

We described this strategy as “specifically referring to the multiple-choice questionnaire options and proceeding by elimination to infer the intended meaning” (cf. [Figure 18](#)). In our interview data, this strategy was always combined with other strategies but we felt it was important to code it as a specific method of inference. It is different from the other strategies in the fact that it does not rely on the videos per se, but rather is more of a test-taking strategy that was the result of our test design choices. Indeed, following Taguchi et al.’s methodology, after watching the videos, participants would have to demonstrate their understanding of the target utterance by selecting an answer from a set of four options (c.f. [sub-section 5.2.1](#)). We will discuss the impact of this strategy in [Chapter 9](#), but as the samples we provided in [Figure 18](#) illustrate, this process of elimination relied on participants looking at the four multi-choice options and narrowing down their answer based on whether or not some of the options seemed possible. We coded for this strategy any time participants clearly expressed that they rejected other answers as part of their inferring process. It is however very likely that this elimination process was part of the participants’ overall process even though they may not have explicitly verbalized it. Taguchi et al. did not mention this type of strategy (either as a category or example), and it is possible they included any such examples in the logical reasoning category as it does indeed involve a deductive process. However, we thought it was important to separate the two categories since the process of elimination strategy was only made possible because of the way our instrument was designed. Contrary to all the other types of strategies, this one is not a strategy that would normally be used in the comprehension of French indirect or direct meaning in a naturalistic setting.

### 8.5. Data analysis

As previously discussed, we only coded the transcriptions of the participants’ retrospective verbal reports for incidence of inferencing strategies that learners used to interpret indirect meaning (we did not code the verbal reports on literal meaning). While Taguchi et al. did not explicitly discuss how they determined the inferencing strategy frequency in their 2016 paper, we assumed<sup>59</sup> they followed the same protocol Taguchi had used in her 2002 paper: “A discrete strategy was counted only once per item, even if it appeared repeatedly within each item”

---

<sup>59</sup> Taguchi et al. mentioned using Taguchi (2002)’s strategies as the base for their coding so we assumed they used the same protocol (p. 689): “A total of 72 verbal reports [...] were coded based on Taguchi (2002) for instances of strategies...”

(Taguchi, 2002, p. 161). Since we wanted to be able to compare our results with Taguchi et al.'s as part of our replication, we followed the same system. However, as mentioned in the previous section, our approach was data-driven, which explained why we ended up with additional strategies that Taguchi et al. did not find. Since we also wanted to investigate how successful these strategies were at reaching the correct interpretation, we made note of the following:

- whether participants had selected the correct MCQ option during the in-class test.
- whether participants expressed understanding the target utterance on a sentence-level meaning (either referencing the literal meaning of target sentence in French or translating it).
- whether participants selected the correct MCQ option or not.

Regarding this last point, we should note that, because our verbal reports were reflective of participants' inferencing process, participants would sometimes change their answer as they were verbalizing their thought process. [Excerpt 6](#) illustrates this phenomenon.

### Excerpt 6

*Video Q40 - OPI 8 - Participant 5 [low-scoring]*

**Participant 5:** [...]. So, (3.0) number four would be my immediate thing because they use, would be my immediate answer, because they use the number signifier.

**Researcher:** Mhm

**Participant 5:** But um:: with the other answers and the context, I got number one, because you don't just go to sleep for 15 minutes that is

**Researcher:** [Okay]

**Participant 5:** not] realistic. So usually when you fall asleep in the movie, you're not very entertained because it's, you know, it's dark and also the midnight premiere thing that kind of made you think that was night so maybe she just thought it was really boring?

Participant 5 first selected option #4, which is incorrect and based on a literal interpretation of the target utterance—and that was the option they had selected in the online survey—based on a keyword inferencing. However, as he was thinking about his answer, he then switched to option #1 (correct interpretation) after looking at all the MCQ options (process of elimination). He used logical reasoning and background knowledge to move past the literal interpretation, as he realized that one does not just sleep for 15 minutes when watching a movie. Since the participant's final choice was the correct MCQ option we noted this verbal report as reaching accuracy.



Earlier, we explained that there were times when we felt that we departed from the initial interview protocol by providing additional information to the participants (for example, when participants were unable to understand the dialogue at all, we translated parts of it to see if they would be able to access the intended indirect meaning once they had access to the literal meaning). Strategies that derived from this “off-script” part of the interviews were not coded or included in the overall count or strategy distribution numbers.

## Chapter 9 - Analyses of Retrospective Verbal Interviews: Results and Discussion

This chapter will be dedicated to analyzing and discussing the results from our interview data. We will follow the same organization as Taguchi et al. (2016). First, we will look at potential differences between the low-scoring and the high-scoring groups ([section 9.1](#)). Based on a comparison of accuracy rates and access to the sentence-level meaning of the target utterances, we will identify noticeable patterns of strategy use by low-scoring and high-scoring participants. Each sub-section will include a discussion of notable patterns and a comparison with Taguchi et al.'s findings. Excerpts from our interview transcripts will be used to illustrate differences between the two participant groups, and whether the strategies they used were successful or not. In the second section ([section 9.2](#)) we will look at the impact of indirect meaning types (indirect refusals, irony, indirect opinions) on strategy use. Again, numerical data will be presented to reflect on the frequency of use for each category of meaning, and compare it with Taguchi et al. We will then discuss noteworthy examples of strategy use that are specific to each meaning type. Of course, we expect that factors like participants' proficiency and category of items will interact and we anticipate seeing an impact of both participant group and meaning type on the use of strategies, based on the literature we reviewed ([section 3.3](#)).

### 9.1. Inferencing strategy use and effectiveness among high- and low-scoring participants

Two major differences between the low and high-scoring participants were brought to light through our analyses of the interview data. The first one, which we expected, was about the participants' accuracy. The second was about was the type and quantity of strategies each group seemed to favor. Combining these two types of results allowed us to get better insights into which strategies were more efficient or successful.

#### *9.1.1. Access to intended meaning vs. access to literal meaning of the target utterance*

Participants were considered as having comprehended the intended meaning if they were able to select the correct MCQ option. [Table 21](#) provides the accuracy rate (percentage of times the

participants selected the correct MCQ option as their final answer<sup>60</sup>) for the nine indirect meaning items targeted during the retrospective interviews, as obtained during the in-class test, and then as obtained during the retrospective interviews to show the discrepancies.

**Table 21**

*Accuracy rate on interview items per group*

	In-class test	Interview
	Indirect items ( $n = 9$ )	Indirect items ( $n = 9$ )
High scoring	91.67%	88.89%
Low scoring	30.56%	47.22%

As [Table 21](#) shows, during the in-class test, the accuracy rate on the nine indirect meaning items selected for the interviews was much higher for the four high-scoring participants (91.67%) than for the low-scoring group (30.56%). A similar difference was observed during the interviews: the high-scoring group was accurate 88.89% of the time while the low-scoring group only selected the correct MCQ option 47.22% of the time. It is interesting to note that the difference between the two groups decreased during the interview: the high-scoring participants scored slightly lower during the interview than they did during the in-class test, but the low-scoring participants scored appreciably higher. Sometimes, participants selected a different MCQ option than they did on the in-class test (and admitted they did not remember what they had answered on the test), but as we discussed in [section 8.5](#), the MCQ selection process during the interview was also rather fluid: participants would sometimes change their mind as a result of having to explain their thought process.

As a point of comparison, Taguchi et al. (2016) reported that their high-scoring students made correct inferences 97% of the time, while the success rate was 70% in the low-scoring group. Our participants (especially our low-scoring group) had a much lower accuracy rate than Taguchi et al., which is consistent with what we had previously discussed regarding the impact of proficiency on comprehension, and the fact that our participants were likely less proficient overall than Taguchi et al.’s participants.

<sup>60</sup> Not including “off-script” segments of the verbal reports as previously mentioned.

While a debate around the obligatory processing of literal meaning for indirect meaning process is still ongoing (as discussed in [sub-section 2.6.2](#)), we know that the inability to understand the meaning of an utterance at the sentence level negatively impacts one's ability to understand its illocutionary force and reach the implicature (cf. Field, 1998, 2008; Graham, 2017; Rost & Ross, 1991; Zoghلامي & Hilton, 2021; as well as section [2.1](#) and [2.2](#)). For this reason, proficiency is often regarded as a determining factor for L2 comprehension success (of indirect and direct meaning) as explained in [section 3.1](#) (cf. The Listening Threshold Hypothesis, Vandergrift, 2006). One thing that transpired during our interviews was the fact that our low-scoring participants often struggled with the linguistic input. In fact, we only noted two instances for which the low-scoring participants were able to understand the target utterance (at least on a sentence-meaning level<sup>61</sup>) of the indirect meaning items (5.56%). Indeed, low-scoring participants frequently hesitated in their answers, and admitted to not understanding part or most of the dialogue, including the last comment, which was the focus of our study. Because they did not understand what was being said, accessing the intended meaning became a lot more challenging. It is interesting to see however, that while this low-scoring group only accessed the target sentence-level meaning 5.56% of the time, they were still able to select the correct MCQ option 47.22% of the time (cf. [Table 21](#)). That means that there were instances when they were able to use some inferencing strategies to remedy their lack of sentence-level comprehension and compensate for the fact that they could not access the full dialogue. This is consistent with what Taguchi (2002) reported about lower-proficiency learners' ability to comprehend indirect meaning despite their inability to access sentence-level meaning: "The verbal protocols clearly document that the learners were able to interpret implicit messages even when they did not completely understand the language in the test sentences" (p. 164). Admittedly, there were also instances when the participants simply guessed the correct MCQ option, which means that the accuracy rate is probably inflated.

On the other hand, during the interviews, the high-scoring participants were able to understand the full dialogue (or at least the target sentences on which the indirect meaning was based) 75% of the time. Every time they had access to the target sentence-level meaning, they were able to reach the intended meaning without any issues. We did not observe Thomas' (1983) "level 2 misunderstandings" (which Bell & Attardo, 2010, p. 432 refer to as "Level 3: failure to understand the pragmatic force of utterances"): when the literal meaning of the utterance is

---

<sup>61</sup> what we mean by "sentence-level meaning" here is what Thomas (1995) refers to as the first level of meaning (cf. section 2.1): what is said.

processed correctly but the hearer is not able to recognize the intended meaning. In those instances when the high-scoring participants admitted not being able to fully hear the target sentence (nine out of 36 indirect meaning verbal reports) they were sometimes able to understand enough of it to make out the rest of the intended meaning, or they used contextual cues to select the correct MCQ options. There were only four instances (out of 36 verbal reports) for which a high-scoring participant was unsuccessful in selecting the correct MCQ option.

We will now turn our attention to the frequency of strategies used by each group, to identify differences in use between low and high-scoring participants, and compare these findings with Taguchi et al.'s.

### 9.1.2. Differences in strategy use by participant group

[Table 22](#) shows the frequency of inferencing strategies used by our high-scoring and low-scoring participants for the indirect meaning items.

**Table 22**

*Frequency of Inferencing Strategies by Participant Group*

Strategy	Low-scoring learners	High-scoring learners
Auditory cues	6 (7.2%)	9 (12.7%)
Visual cues	19 (22.9%)	13 (18.3%)
Background knowledge	6 (7.2%)	7 (9.9%)
Keyword inferencing	27 (32.5%)	5 (7%)
Logical reasoning	6 (7.2%)	29 (40.8%)
Other:		
Setting/Situation	4 (4.8%)	2 (2.8%)
Familiarity with characters	0	2 (2.8%)
Adjacency pair rule	1 (1.2%)	3 (4.2%)
Process of elimination	14 (16.9%)	1 (1.4%)
Total	83	71

We identified a total of 154 instances of strategy use across the 72 verbal reports focusing on indirect meaning comprehension. Comparatively, Taguchi et al. (2016) only identified a total

of 100 strategies across their 72 verbal reports (45 strategies in their low-scoring group and 55 strategies in the high-scoring group). Of course, it is possible that this discrepancy could be the result of variations in our coding systems.

Another striking difference between Taguchi et al. and our results is the fact that their low-scoring participants used fewer strategies than their high-scoring participants, but we observed the opposite pattern. Our low-scoring group relied on a greater number of strategies than the high-scoring group (83 compared to 71). We believe the reason for this is that our low-scoring participants had to use strategies to make up for their lack of linguistic comprehension ability. Where high-scoring participants were able to understand the target sentence easily and derive the intended meaning from it, low-scoring participants needed to find additional sources to interpret the dialogue. To further investigate this discrepancy, we will now analyze and discuss each inferencing strategy and contrast its use by the low-scoring and high-scoring groups. Each sub-section will also include a comparison with Taguchi et al.'s findings. We will start with the inferencing strategies that, in our opinion, presented the most noteworthy patterns of use by one or both groups of participants.

#### 9.1.2.1. Keyword inferencing

As shown in [Table 22](#), the keyword inferencing strategy was used predominantly by low-scoring participants who used it more frequently than any other strategies (it appeared in 27 out of 36 verbal reports). High-scoring participants only relied on keywords in five of the 36 verbal reports. Overall this strategy led to the selection of the correct MCQ option 31.25% of the time (29.63% for the low-scoring group and 40% for the high-scoring group), which means that it was not a very effective inferencing strategy.

As mentioned earlier, during our interviews, our low-scoring participants often admitted to not fully understanding the dialogues. Therefore, they had to rely on the few keywords/lexical items they did hear (or thought they heard) to try to understand the dialogue. Sometimes these keywords were clearly identified lexical items, sometimes they were expressions or lexical chunks. Additionally, keyword inferencing was also used in combination with the elimination process strategy (which we will discuss in [sub-section 9.1.2.3](#)) as participants would discard some of the multiple-choice options provided in the test based on the fact that they reported not having heard a specific keyword contained in the multi-choice answers.

[Excerpt 7](#) illustrates a keyword strategy that resulted in the wrong interpretation. Participant 7 (low-scoring) was focusing on the following indirect refusal item:

**IRF 2** - *This is an interaction between two friends, Fabrice (male speaker) and Marie (female speaker). They are in a cafeteria and they are talking about having something to eat.*

**Dialogue:**

Fabrice: Salut Marie! Je t'ai commandé quelque chose à boire. J'ai pris un chocolat chaud; il fait si froid aujourd'hui.

Marie: Merci Fabrice!

Fabrice: Est-ce que tu voudrais quelque chose à manger? Je n'ai rien commandé car je ne savais pas ce que tu préfères.

Marie: C'est bon. J'ai déjà déjeuné à la maison.

**Multiple-choice options :**

1. Marie wants something to eat together with the hot chocolate.
2. Marie will have her hot chocolate at home.
3. Marie wants to eat a piece of apple pie together with the hot chocolate.
4. Marie doesn't want anything to eat.

Participant 7 picked up on the keyword for “hot chocolate,” which appears in three of the four MCQ options. However, she also heard the word for “house,” which led her to select option #2.

**Excerpt 7**

*Video Q2 – IRF2 - Participant 7 [low-scoring]*

**Researcher:** So these are your options. What are you thinking?

**Participant 7:** Um, I'd say two

**Researcher:** Okay.

**Participant 7:** I don't remember what I put but based off of what I just watched two.

**Researcher:** And that is that is what you put in class. So what do you, how did you get to that answer? Tell me a little bit about that.

**Participant 7:** Um, **so all of the answers have to do with...** Well, **first you could hear them say hot chocolate at some point. I remember them saying “chocolat” and like “chaud.”**

**Researcher:** Mhmm

**Participant 7:** And **then the last word he says is “maison,” which is “house”**

**Researcher:** Mhmm

**Participant 7:** So that's, yeah, that's kind of how I put it together.

[Excerpt 7](#) clearly shows how Participant 7 tuned in to a word that was in one of the MCQ options. We may remember that option #2 here is based on Taguchi et al.'s (2016) Principle 2 (“The option contains words taken from the last part of the conversation,” p. 687). It is therefore

not surprising that a participant who relies on keyword inferencing (due to their inability to access the full dialogue meaning) would select that option. Overall, keyword inferencing strategies alone proved to be fairly unreliable. Only when participants combined it with other strategies were they able to more consistently select the correct MCQ option (though again we should keep in mind that overall accuracy rate for this low-scoring group was only 47.22% and the success rate for this strategy was about 30%). Even then, the deciding factor that led to the selection of the correct MCQ would usually stem from a strategy other than keyword inferencing, as illustrated in [Excerpt 8](#).

### Excerpt 8

*Video Q6 – IRF6 - Participant 3 [low-scoring]*

**Researcher:** Alright, so the options is number one. Mr. Bodnar and Stefanie are not going to meet on Tuesday. Number two, Mr. Bodnar and Stephanie are both available on Tuesday. Three, Mr. Bodnar teaches at University on Tuesdays and four Stephanie has no doubt she can turn in her homework next class. What do you think?

**Participant 3:** Oh gosh, okay, um (2.0). Is it that he teaches at the University on Tuesdays?

**Researcher:** Yeah, I think that's consistent with what you had answered. So what, why did you decide to answer that one?

**Participant 3:** Oh, *I heard Tuesday. I heard Tuesday.* Um: It is it number one where they can't meet on Tuesday, *because their facial expressions make me think that like they were trying but they cannot meet.*

**Researcher:** Um, yes.

**Participant 3:** I thought it was because he is teaching on Tuesday. And that's why they couldn't meet but...

**Researcher:** Yeah. So yeah, because you know if you said the word Tuesday. It actually shows up in three of the answers.

**Participant 3:** Yeah, yeah.

In [Excerpt 8](#), we see Participant 3 first selecting the MCQ option associated with Principle 2 (it contains words from the last part of the dialogue) based on a keyword she heard. She mentioned that she heard the word for Tuesday but that word is also part of two other MCQ options. She then switched her answer to the correct MCQ option, based on both characters' facial expressions (visual cues will be discussed next).

By contrast, the high-scoring group made very limited use of keyword inferencing. When they did (five out of 36 verbal reports) it was usually because they were not able to understand the full dialogue or the target sentence (and all five instances were in the indirect opinions



category). In three of five cases they failed to select the correct MCQ option, likely because they only relied on keyword inferencing (the other two times they combined it with other cues—visual cue for one, and a mix of visual, adjacency and familiarity with character for the other) to select their answer, (as illustrated by [Excerpt 9](#)).

### **Excerpt 9**

*Video Q39 – OPI7 - Participant 33 [high-scoring]*

**Researcher:** Alright, so based on her last comment.

**Participant 33:** I think she liked the massage, she said something about er “my back was never something before the massage.”

**Researcher:** Okay.

**Participant 33:** I didn't quite understand all of it, but it sounded like she was speaking about it positively.

**Researcher:** I think... let me see... When we did it in class, you picked number Two “Sophia believes the massage for the back was not good.”

**Participant 33:** Okay. I think that was just one that I didn't understand very well either time

In [Excerpt 9](#), Participant 33 provided an incorrect answer here (opposite of the intended meaning and a different answer than what he selected in class). He explained that he could not understand the last sentence completely and of course the one word he did not get (the fact that the character said her back did not hurt before the massage, implying the massage was not good because her back hurt now) changes the meaning of the whole sentence. It is not clear whether he was referring to visual or auditory cue when he mentioned that he thought the character was “speaking about it positively” but in the end, it appears that it was his reliance on a partial sentence (as opposed to the whole sentence) that prevented him from accessing the intended meaning.

This difference in use of keyword inferencing between low-scoring and high-scoring participants we observed was also one that Taguchi et al. noted: “The low-scoring group used key word inferencing strategies more often than the high-scoring group” (p. 690) and they too attributed it to lower proficiency: “Previous findings have shown that less-proficient listeners tend to go directly to the first association of the key word and the referent” (p. 687). Taguchi (2002) had similar findings and references Ross (1997) for evidence of the fact that “the hearing of a key word is the most common processing level for lower proficiency L2 listeners” (Taguchi, 2002, p. 167). It seems that this preference from lower-proficiency learners to use

this type of strategy is not specific to indirect meaning inferencing but more of a general comprehension strategy that reflects their word-by-word processing (Rost & Ross, 1991). We also note an overall higher number of keyword inferencing strategies in our data as compared to Taguchi et al. (2016). This is likely both a consequence and confirmation of the fact that our participants had a lower proficiency level than Taguchi et al.'s participants.

Given the preponderance (27 out of 36 verbal reports included keyword inferencing) and unreliability of keyword inferencing among the low-scoring group, it is not surprising that their performance on the test was low. These results were congruent with Taguchi et al. (2016), who also found a higher use of keyword inferencing strategies by low-scoring participants which resulted in “unsuccessful guessing” (p. 695). This is also consistent with Zoghلامي’s (2016) findings on L2 listening comprehension: her experimental study showed that unskilled listeners failed to activate the lexical meaning of the oral input and were not able to understand an audio recording. While her lower-proficiency learners did attempt to a mix of top-down (use of context) and bottom-up (linguistic knowledge) strategies to infer meaning, they were not always successful.

#### 9.1.2.2. Logical reasoning

The pattern of use for the logical reasoning strategy was the striking opposite of the keyword inferencing: high-scoring participants used logical reasoning more than any other strategies and at a much higher rate than low-scoring participants. Twenty-nine of the 36 verbal reports from high-scoring participants included logical reasoning whereas low-scoring participants only used it six times (cf. [Table 22](#)). As discussed in [sub-section 9.1.1](#), the high-scoring participants’ ability to fully understand the dialogue and access the sentence-level meaning of the target utterance was likely the primary factor for their frequent use of logical reasoning since this strategy, as defined by Taguchi (2002), was based on the ability to recognize the literal meaning and deductively work towards the intended meaning. This also explains why low-scoring participants, who rarely understood the target sentence, were not able to use the logical reasoning strategy as often as high-scoring participants, and had to settle for keyword inferencing instead. Yet, when the low-scoring participants did use logical reasoning (six out of 36 verbal reports), it always resulted in the selection of the correct MCQ option. Just like with the low-scoring participants, logical reasoning helped high-scoring participants select the correct MCQ option 100% of the time. The main reason why logical reasoning was so

successful, and one of the strategies predominantly used by high-scoring participants is because it relied on a more complete understanding of the full dialogue; and we have shown in Chapter 3 ([section 3.1](#)) that understanding the linguistic meaning of an utterance can be crucial to understanding its pragmatic meaning. These results are therefore not surprising.

When low-scoring participants used logical reasoning, they always combined it with other strategies. High-scoring participants also used logical reasoning strategy in combination with other strategies (20 times) but there were nine instances in which they only used logical reasoning to reach the correct interpretation and did not attend to any other contextual cues (or at least did not verbalize doing so),<sup>62</sup> as illustrated in [Excerpt 10](#) (used previously in [sub-section 8.4.4](#)).

### Excerpt 10

*Video Q9 – IRF9 - Participant 28 [high-scoring]*

**Researcher:** OK, so here are the options [showing the MCQ on screen], what do you think?

**Participant 28:** That she will not try tomorrow with Eric at 7pm.

**Researcher:** Okay. Perfect. So how did you get to that answer?

**Participant 28:** Again, the dialogue, just **the fact that she has a concert to get to at 7:15 so obviously she wouldn't be able to make a dinner at seven o'clock.**

**Researcher:** Very good. Any other things that you attended to

**Participant 28:** No just the dialogue, that's it.

**Researcher:** Okay. Thanks.

When asked if there were other things that helped him, he said he just based his answer on the dialogue (see underlined sentence). The same participant, for the previous video, had actually used visual cues in addition to logical reasoning, as shown in [Excerpt 11](#).

### Excerpt 11

*Video Q6 – IRF6 - Participant 28 [high-scoring]*

**Researcher:** Alright, so here are four options for this one—can you see them on the screen? you had 1) Mr. Bodnar and Stefanie are not going to meet on Tuesday, 2) Mr. Bodnar and Stephanie are both available on Tuesdays, 3) Mr. Bodnar teaches at the University on Tuesdays, 4) Stephanie has no doubt that she will be able to turn in her homework to Mr. Bodnar next class.

**Participant 28:** uh they will not be meeting on Tuesday.

**Researcher:** Okay perfect yes, that was the correct answer. So same thing. What were you thinking when you chose the answer, what helped you decide on that answer?

---

<sup>62</sup> This was particularly true for indirect refusals items so we will discuss this pattern in section 9.2

**Participant 28:** Mmh, again, mm a little *the dialogue, because he was saying that he's not at the University on Tuesday.* Um, I guess, also her facial expression whenever he said that she made like a like. er a face. Yeah.  
**Researcher:** Very good.

As exemplified by [Excerpt 11](#), when logical reasoning was used in combination with other strategies (such as visual or auditory cues), the latter were often used as secondary/supportive strategies that supported the main reasoning (cf. Yus' supportive sources in Yus, 2000—cf. [sub-section 2.6.4](#)).<sup>63</sup> Because the literal meaning was more salient to participants who used logical reasoning (primarily the high-scoring participants), use of other strategies often came second. This is also consistent with what we know from Giora's Graded Salience Hypothesis (cf. [section 2.6](#)). Since the literal meaning of the target utterance was easily accessible to high-scoring participants, they used logical reasoning primarily to understand indirect meaning.

Our findings regarding logical reasoning use by high-scoring participants are in line with Taguchi et al. (2016) who also found that their “high-scoring students used logical reasoning more often than their low-scoring counterparts” (p. 690). They also observed that logical reasoning was a particularly effective strategy, but contrary to our data, their low-scoring participants used logical reasoning strategies as often as they used keyword inferencing strategies. In contrast, our low-scoring participants used keyword inferencing at an overwhelming higher rate than logical reasoning. As we previously mentioned though, we attribute this difference to the fact that our low-scoring participants were likely less proficient than Taguchi et al.'s participants (based on proficiency self-assessment results—cf. [sub-section 7.1.2](#)). The fact that Taguchi et al. found a total of 49 instances of logical reasoning in their 72 reports but we only identified 35 total, is further evidence of our participants' lower proficiency level.

### 9.1.2.3. Visual cues

Visual cue strategy was the second most used strategy by both low-scoring (19 out of 36 reports) and high-scoring participants (13 out of 36 reports). This strategy led low-scoring participants to the correct interpretation about 52.6% of the time, whereas high-scoring participants used it successfully 100% of the time. A closer look at the strategy combination patterns for each

---

<sup>63</sup> This was particularly true for Irony items and we will discuss this use of combined strategies in section 9.2

participant group showed that the visual cues were often used by the high-scoring group in combination with logical reasoning (see previous sub-section)—whereas low-scoring participants would usually use visual cues along with keyword strategies. Knowing that logical reasoning led to more accurate answers than keyword strategies, we believe that the use of these combined patterns was what impacted the success rate of the visual cue strategies for the two participant groups. Only twice was this visual cue strategy used by itself (without combining it with other strategies) and it was by low-scoring participants, who were actually able to select the correct MCQ option (though it is plausible that having four MCQ options also helped). In both cases the character’s facial expression in the video was considered negative by the participants and this paralinguistic source was salient enough for the participants to derive their interpretation, as illustrated in [Excerpt 12](#).

### Excerpt 12

*Video Q40 – OPI8 - Participant 3 [low-scoring]*

**Researcher:** Again. Okay, so they're talking about a movie that Florence saw at the movie theater. [shows MCQ options on the screen]  
(3.0)

**Participant 3:** Um:::. Is it that she thought the movie was boring?

**Researcher:** Okay, yeah. Yeah, that was also what you answered. So what made you get to that answer?  
(3.0)

**Participant 3:** I don't know. I feel like it's... I don't know. It's like, where *she just seems like kind of annoyed [makes facial expressions that mimics the character in the video]*. And so I think that she's saying that she didn't like it very much.

**Researcher:** Okay, so basically, the way that she looks right?

**Participant 3:** Yeah.

Later in this interview, Participant 3 asked what the target sentence said because she did not understand it at all. Yet, the character’s facial expression was sufficiently clear and accessible for her to guess the correct MCQ option.

In some instances, non-verbal cues helped participants get to the correct answer (as illustrated in [Excerpts 8](#), [11](#), and [12](#)) but there were also times when low-scoring participants misinterpreted the visual cues (as in [Excerpt 13](#)), which then resulted in them selecting an incorrect answer (often opposite to the correct answer in meaning). In [Excerpt 13](#), Participant 7 misinterpreted the character’s body language, which led him to select an incorrect answer (the opposite of what the intended meaning was).

### Excerpt 13

*Video Q39 - OPI 7 - Participant 7 [low-scoring]*

**Participant 7:** I'd say three. Yeah. You're still muted.

**Researcher:** Oops. Alright, so yeah, so three. Tell me how you got to that answer.

**Participant 7:** Um, well, it starts in the question itself, it says that Sophia got the message

**Researcher:** Mhmm.

**Participant 7:** But not Gauthier, so that automatically takes out one and it seems like he's asking her about it.

**Researcher:** Mhmm.

**Participant 7:** You do hear her say her back

**Researcher:** Mhmm.

**Participant 7:** Just like her body language seems like she liked it.

**Researcher:** Yeah, she does say the word for “back” but I mean that's in answer number three, two and four.

**Participant 7:** Yeah, it's not helpful, but (3.0)

**Researcher:** Yeah, but based on her body language, it looks like she enjoyed.

**Participant 7:** Yeah.

We can see four different strategies at play in [Excerpt 13](#). Participant 7 selected option #3 (“Sophia liked the massage she got on her back,” which was the opposite of the intended meaning) following this process that he verbalized: First, he eliminated one answer based on the video description (what we believe he refers to as the “question” in the verbal report).<sup>64</sup> Next, he turned his attention to a specific keyword he heard in the video. He then mentioned the fact that the character’s body language made him think that she enjoyed her massage. The researcher pointed out that the keyword “back” was actually part of the three remaining options, which would not help in the selection process. At the end, Participant 7 confirmed that he chose his answer (which was incorrect) based on visual cues—not on this specific keyword. Of course, it is unclear why Participant 7 interpreted this specific body language as a positive feeling. In the video, the actress is playing the role of a young lady whose back is hurting after she received a massage and when asked about it, she points to her back, but also makes facial expressions (raising eyebrows, eye widening, no smile) that should have been associated with negative feelings. As mentioned in [Chapter 5 \(sub-section 5.1.5\)](#), our two actors were not professionally-trained, and while we asked them to pay attention to their body language,<sup>65</sup> it is also possible that their facial expressions may not have properly reflected a certain intended meaning.

---

<sup>64</sup> We will discuss these strategies in two upcoming sub-sections.

<sup>65</sup> As a reminder, the actors “were briefed on the general purpose of the study. They were instructed to play the scene in a way that was as natural and authentic as possible, but to pay attention to gesture, intonation and facial expressions. However, we did not give them any instructions as to which specific gesture, intonation or expressions to use.”

However, we know that even in real life, body language is not always easy to interpret, even for native speakers.

Our data shows that both low-scoring and high-scoring participants made use and benefitted from the multimedia input and visual component that the video format of our test offered. By comparison, Taguchi et al. also reported their participants using visual cue strategies but not to such a high level. One of the reasons we believe our low-scoring participants used it more frequently (more than Taguchi et al.'s participants and more than our high-scoring participants) is because this strategy was used to compensate for their lack of linguistic ability. Indeed, to remedy their inability to comprehend the full dialogue, they often had to rely on other clues in the video, and some visual cues were salient enough for the participants to derive meaning (though not always correctly). And while the success rate of these visual cue strategies was still limited for low-scoring learners (52.6% as previously noted) this is still a considerably powerful inferencing tool considering these participants only understood the target sentence 5.56% of the time.

#### 9.1.2.4. Process of elimination

The process of elimination strategy was mostly used by low-scoring participants. One high-scoring participant relied on the MCQ options in one verbal report because she was questioning if she had heard the target sentence correctly and was considering two options (but ended up selecting the correct one). Process of elimination appeared in 14 of the 36 verbal reports on indirect meaning comprehension by low-scoring participants. This was their third most frequently used strategy category. As mentioned previously ([sub-section 8.4.5.4](#)), it is very likely that participants actually relied on the MCQ options as part of their selection process at a much higher rate but we only coded for instances in which the participants specifically verbalized rejecting certain MCQ options. It appears that this strategy was used predominantly by low-scoring participants as a way to compensate for their lack of linguistic comprehension ability. Because they could not access the full meaning of the dialogue, they often had to use this test-taking strategy to decide which option they should choose. We think this was somewhat of a design flaw: being able to select the correct MCQ option does not necessarily equate to being able to interpret indirect meaning (without relying on the four MCQ options). Out of the 14 verbal reports from low-scoring participants that contained a process of elimination strategy, only half led to the selection of the correct MCQ option. Just like with the keyword inferencing

and visual cues, this strategy was not very effective in reaching the correct interpretation as it seemed to rely more on guessing than inferencing. In many cases, participants were able to reject some of the incorrect MCQ options, but eventually selected the wrong one (as we had seen previously with [Excerpt 13](#)).

The process of elimination strategy was always combined with other strategies (sometimes, up to five different strategies, e.g., process of elimination, visual cue, keyword inferencing, background and logical). Similar to the way the rest of the inferencing strategies relied on contextual sources (cf. Yus 2000) to reach the intended meaning, participants turned to the contextual sources to reject MCQ options and infer meaning. The most common combination was with keyword inferencing (as we have seen in several excerpts already—and probably due to the fact that this was used primarily by low-scoring participants who used keyword inferencing at a very high rate) as show in [Excerpt 14](#). This excerpt is based on the following video:

**IRF 9 Q9** *This is a conversation between two friends, Eric (male speaker) and Estelle (female speaker). They are talking about making plans to go to China King to try their all-you-can-eat buffet.*

**Dialogue:**

Éric: Quand est-ce qu'on va à China King pour essayer leur buffet à volonté?

Estelle: Quand tu veux Éric.

Éric: Demain à 19h, ça te va?

Estelle: Ben, j'ai des billets pour aller au concert de Drake à 19h15.

**Multiple-choice options :**

1. Estelle is going to a concert this afternoon.
2. Estelle won't try China King's buffet with Eric tomorrow at 7PM.
3. Estelle and Eric will go to China King together tomorrow at 7PM.
4. Estelle does not like all-you-can-eat buffets.

**Excerpt 14**

*Video Q9 – IRF9 - Participant 5 [low-scoring]*

**Participant 5:** Um (7.0) [sighs]. I think it's two.

**Researcher:** Tell me how you got that?

**Participant 5:** The reason I picked that again, **there wasn't really a signifier of affirmation**. Um, and also **I kind of got stuck in the last few words which, to my interpretation were, it was two numbers “dix et quinze” and there was a word between them, but I couldn't pick it up at something like “nouveau” but that probably isn't correct. Um, so I assume because it was two numbers that it might be a time of some sort.** So maybe she was explaining that she had other things to do at that time. **And that's further strengthened by the fact that the first answer is,**



she's going to a concert this afternoon. Which may have nothing to do with the video at all. But that's just me inferencing based on the answers because when most students are taking exams or tests or doing homework, sometimes the questions, the process, we're taught the process of elimination in elementary school like, you'll see a couple of answers that you can immediately rule out and some answers can have context to others. So that's the strategy that I was kind of using for that one. But that's what I got from that particular video.

**Researcher:** Yeah, and it's totally fine.

In the verbal report, Participant 5 selected option #2 which was the correct answer. He picked up a few keywords which he associated with a time. A time appeared in two MCQ options (#2, the correct interpretation of the intended meaning, and #3 the opposite of the intended meaning). Using this keyword strategy, he guessed that the character might be referring to something she had to do, but he also noticed that option #1 was referencing the character going to a concert. Based on this specific MCQ option, he figured that the dialogue likely alluded to the character having other plans (possibly going to a concert): “So maybe she was explaining that she had other things to do at that time. And that's further strengthened by the fact that the first answer is, she's going to a concert this afternoon.” So, while he did not understand the target sentence at all, he not only proceeded by elimination, but also used information from the MCQ resourcefully to make inferences about the type of information that was in the dialogue. We should note here that Participant 7 was responsible for eight of the 14 occurrences of elimination process strategy among the four low-scoring participants, and as he recognized in [Excerpt 14](#), was particularly aware of this strategy process.

#### 9.1.2.5. Auditory cues

The difference in the frequency of use of auditory cue strategies between low-scoring and high-scoring participants was not very large. Out of the 36 verbal reports for each group, low-scoring participants used auditory cues six times and high-scoring participants nine times. This was the third most commonly used strategies among the high-scoring group. What is interesting however is the high success rate in both groups: five out of the six times the low-scoring group used auditory cues, they ended up selecting the correct MCQ option, and in the case of high-participants, this strategy helped them reach the correct answer 100% of the time. For the high-scoring participants, we noted a clear combination pattern: auditory cues were always combined with logical reasoning (and in a few instances with additional strategies such as visual cues, background or setting) and often served as secondary/supportive cues (just like visual cues), as shown in [Excerpt 15](#) below (which focuses on an indirect opinion).

## Excerpt 15

Video Q34 – OPI2 - Participant 33 [high-scoring]

**Researcher:** Alright, so based on that last statement.

**Participant 33:** She did not like the dress very much because she said it felt like it was in the 80s.

**Researcher:** Okay. Yeah. Was there anything else that you picked up on?

**Participant 33:** Her sarcastic tone.

No noticeable combination pattern emerged from the use of auditory cues by the low-scoring group. However, it seems that while low-scoring participants used auditory cues less frequently than visual cues, the former was actually more effective in reaching the correct interpretation. Of course, we should refrain from generalizing given the small number of occurrences of this strategy in the low-scoring group, but the fact that participants used both visual and auditory cues to derive meaning from sentences they did not understand in the first place is an important aspect of the L2 comprehension process. Along the same line, Taguchi comments on her participants' use of auditory strategies (2002, p. 165):

The results show that the paralinguistic features served as salient cues in understanding the attitudinal/emotional aspects of a message, and helped learners to draw inferences about the speaker's intended meaning, particularly when they could not comprehend the language.

Just like with visual cues however, one cannot be completely sure that they are interpreting auditory cues correctly. Even in real life, intonation and tone of voice can sometimes be misinterpreted, by native speakers and L2 learners alike. If on top of it, one is not able to understand the literal meaning of what is said, it can become even more challenging, as shown in our next excerpt. In [Excerpt 16](#), Participant 7 (low-scoring) really struggled with the dialogue (which was hard for many participants—we will discuss this more in depth in [sub-section 9.2.2.1](#)) and could only really access the information available in the scene description as well as part of the dialogue. At the end of the video, the male character states that he has not heard back about his college applications but he is sure he will get admitted because he is an excellent student, to which the female character replies “*Oui, c'est ça, et moi je suis Harry Potter.*” The target sentence was meant to be interpreted as a sarcastic comment from the female character who does not think the male character is an excellent student. However, Participant 7 did not pick up on any of the visual and auditory cues present in the video.

## Excerpt 16

*Video Q24 – IRO8 - Participant 7 [low-scoring]*

**Participant 7:** I remember it was harder. I want to say four, I don't remember what I put it was so long ago.

**Researcher:** Yeah, so you did. You did put four. So how did you get to that?

**Participant 7:** Oh well, I knew, they're talking about like college and like applications and everything like that.

**Researcher:** Mhmm

**Participant 7:** Um, it seemed, I think he asked her if she's done them and she said “yeah I did mine already” and but he hasn't done his right?

**Researcher:** Mhmm

**Participant 7:** And and then, umm, yeah, it doesn't seem like the way she was speaking... it doesn't seem like she was putting him down. It was just like “oh yeah I know you'll probably get to them soon” like, that's what, it kind of seemed to me like the tone and everything.

Participant 7 selected option #4 which was the opposite of the intended meaning (“Iris believes that Paul is an excellent student”). It is impossible to know what in “the way she was speaking” made Participant 7 believe that the female character was being supportive. Maybe this participant is just not skilled at interpreting paralinguistic cues (this was the same participant who misread visual cues in [Excerpt 13](#)).

So far, all the strategies we discussed presented specific patterns of use, either because one group used them more than the other strategies (as it was the case for low-scoring participants and keyword inferencing, or for high-scoring participants and logical reasoning) or because of the differences between the two groups (for example, the process of elimination which was used frequently by low-scoring participants but rarely by high-scoring participants) or a combination of both. We also discussed specific combination patterns (such as logical reasoning and visual/auditory cues by the high-scoring participants) and effectiveness of strategies within each participant group. The rest of the strategies we identified did not show meaningful differences in distribution between the low-scoring and high-scoring group, which is why they were not including in this section (though some will be discussed in our next section—[section 9.2](#)).

To conclude this first section of Chapter 9, we reiterate the notion that the differences between the low-scoring and high-scoring group can be attributed to the participants' (in-) ability to

access the sentence-level (literal) meaning of the dialogue (see [sub-section 9.1.1](#)). Indeed, low-scoring participants admitted on multiple occasions to simply not understanding the words in the video. Since they often had to guess the answer, they were forced to rely on other cues (anything they could see or hear in the video, as well as the provided multiple-choice options from the test questionnaire) to figure out the meaning of the last utterance. This is also supported by the fact that the low-scoring group showed a high frequency of keyword inferencing (since they often could only understand one or two words) and visual cues (as well as process of elimination strategies). Research on L2 listening comprehension (Chamot, 2005; Graham & Macaro, 2008; In'nami & Koizumi, 2022; Mendelsohn, 1995; Roussel & Tricot, 2012; Vandergrift, 2003a, 2003b, 2006; Zhang, 2012, Zoghiami, 2016; Zoghiami & Hilton, 2021) provides many examples of the ways L2 learners (in particular less-proficient learners) can turn to compensation strategies to “overcome any gaps in knowledge of the language” (Oxford, 1990, p. 71) due to their “lack of adequate ‘bottom-up’ information” (Field, 1998, p. 115). Based on our interview data, the same can be said about multi-modal<sup>66</sup> comprehension, with the added benefit of being able to rely on visual cues as another top-down strategy. However, these compensatory strategies—while useful—were only successful about half of the time. As L2 learners become more proficient and have a better understanding of the target language (on a lexical and syntactic level), processing will become more automated (Vandergrift & Goh, 2012), and contextual cues will be used as supportive sources to infer—rather than guess – indirect meaning.

On the other hand, just like Taguchi et al.’s high-scoring learners, our high-scoring participants used logical reasoning at a much higher rate than the low-scoring learners. While these participants sometimes expressed using other strategies (like non-verbal cues), their primary focus was often on making inferences based on noticeable incongruity (whether it was incongruity between what was said and the situation, or what was said and facial expressions, etc.) by following a logical pattern of reasoning, which was only made possible by their access to the sentence-level meaning of the target utterances. As such, they demonstrated a behavior consistent with Sperber and Wilson’s Relevance Theory (1992, 2002—cf. [section 2.5](#)) by detecting the contextual inappropriateness (relying on “multimodal irony markers”—cf. Attardo, 2000a; or Attardo et al., 2013), as well as what Yus (2000) refers to as the “criterion

---

<sup>66</sup> video-assisted listening comprehension

for optimal accessibility” (i.e., multiple activation of various contextual sources and the detection of simultaneous incompatibilities).

In addition to disparities between the two participant groups, our verbal reports also revealed that participants tended to use specific strategies based on the type of indirect meaning they had to interpret. Of course, within each category (indirect refusals, irony, and indirect opinions) we also observed specific patterns of use based on proficiency (as we know the two factors interacted). We will now look at each category of meaning separately to analyze patterns that emerged as a function of indirect meaning type.

## 9.2. Inferencing strategies as a function of item type

In this section, we will first present an overview of the frequency of use per category (indirect refusal, indirect opinion, and irony) and note general differences, before delving into each category more specifically so we can observe and discuss patterns within each category of indirect meaning, and compare with Taguchi et al.’s findings.

[Table 23](#) shows the total number of occurrences of inferencing strategies for each category of items (with percentages within each category).

**Table 23***Frequency of Inferencing Strategies by Item Categories*

Strategy type	Indirect refusals	Irony	Indirect opinions
Auditory cues	5 (11.6%)	9 (14.3%)	1 (2.1%)
Visual cues	6 (14.0%)	13 (20.6%)	13 (27.1%)
Background knowledge	1 (2.3%)	9 (14.3%)	3 (6.3%)
Key word inferencing	10 (23.3%)	7 (11.1%)	15 (31.3%)
Logical reasoning	13 (30.2%)	14 (22.2%)	8 (16.7%)
Other:			
Setting/Situation	1 (2.3%)	4 (6.3%)	1 (2.1%)
Familiarity with characters	0 (0.0%)	1 (1.6%)	1 (2.1%)
Adjacency pair rule	1 (2.3%)	2 (3.2%)	1 (2.1%)
Process of elimination	6 (14.0%)	4 (6.3%)	5 (10.4%)
Total	43	63	48

We notice a much higher number of strategies used for irony as compared to indirect opinions and indirect refusals. Our next table ([Table 24](#)) shows similar numerical data, but breaks down the frequency of strategy use for each group—low-scoring (LS) and high-scoring (HS)—accounting for the potential interaction of proficiency and item type.

**Table 24***Frequency of Inferencing Strategies by Item Categories based on group*

Strategy type	Indirect refusals		Irony		Indirect opinions	
	LS	HS	LS	HS	LS	HS
Auditory cues	3	2	3	6	0	1
Visual cues	4	2	7	6	8	5
Background knowledge	1	0	3	6	2	1
Key word inferencing	10	0	7	0	10	5
Logical reasoning	1	12	4	10	1	7
Other:						
Setting/Situation	1	0	2	2	1	0
Familiarity with characters	0	0	0	1	0	1
Adjacency pair rule	1	0	0	2	0	1
Processing of elimination	6	0	4	0	4	1
Total	27	16	30	33	26	22

[Table 24](#) provides a more nuanced distribution of strategies based on participant group. For example, while indirect refusals had the lowest overall number of strategies, it seems that this was largely due the high-scoring group who used the lowest number of strategies with these items.

We already know that proficiency impacts L2 comprehension and, in the context of the current study, resulted in different strategy use, but we are hoping that our analysis of the strategies by item category might give us some clues as to whether factors such as conventionality (or lack thereof) and access to contextual cues facilitated—or hindered—comprehension.

### *9.2.1. Indirect refusals*

A total of 43 discrete strategies were identified in the 24 verbal reports focusing on indirect refusals, which is the lowest overall number of strategies used for one category. Accuracy rate for the low-scoring group during the interview was 50% for indirect refusals and high-scorers' was 100%.

Logical reasoning and keyword inferencing were the two most frequently used strategies (as shown in [Table 23](#)) for this category of indirect meaning, with a clear distribution between the low-scoring and high-scoring groups: all instances of keyword inferencing strategies came from low-scoring participants and 12 out of the 13 instances of logical reasoning came from the high-scoring group. This is consistent with what we already discussed in the previous section ([section 9.1](#)) regarding strategy preferences for both groups.

Other types of strategies were not used as frequently as logical reasoning and keyword inferencing. Their frequency of use was also lower for indirect refusals than for the other types of meaning (with the exception of process of elimination, which was used at a fairly consistent frequency by low-scoring participant across all types of meaning).

When auditory strategies were used, participants reported picking up on specific phonological cues that were consistent with the illocutionary force of the indirect refusal speech act. For example, Participant 31 (high-scoring) noted a specific intonation when the character indirectly refused an offer to order food (“**Just kind of her intonation, or she also said “c’est bon”** like the

intonation. Like, like, “it's fine.” Like, “I'm good.”—Q2 – IRF2). The same participant also picked up on the character's hesitation in the video Q9 – IRF9 in which the character indirectly refuses an invitation to eat at a restaurant.

As for visual cues, one would expect that certain facial expressions or gestures might be helpful in determining whether someone is indirectly refusing an invitation or offer. It was not necessarily the case in our interview data: half the time, visual cue strategies resulted in the selection of the incorrect MCQ option for indirect refusals. As we discussed in the previous section ([sub-section 9.1.2.3](#)), facial expressions or body language were sometimes misinterpreted by low-scoring participants when trying to understand indirect refusals.

We expected the use of the adjacency pair rule strategy to be much higher for this category, given that this pattern of interaction is particularly relevant in the case of indirect refusals. We think that the almost complete absence of such strategy (only one instance was documented) can be attributed to several factors. First, it is possible the participants noticed the pattern and were primed to expect a certain type of answer (refusal or approval) in the target sentence, but did not verbalize this process. Second, the adjacency pair rule may not have been the most salient cue for participants who then did not rely on or need it. Indeed, the fact that high-scoring participants mostly relied on logical reasoning to successfully infer the intended meaning means they did not need to provide further explaining of their thought process.

What is particularly interesting is the fact that, for this category of meaning more than any of the other categories, high-scoring participants primarily—in eight out of 12 verbal reports—used a single strategy (logical reasoning) to reach the intended meaning. Even when they used multiple strategies (in the other four verbal reports) it was limited to a combination of only two (logical and auditory twice, and logical and visual twice). In contrast, low-scoring participants relied on a greater variety of strategies, sometimes in combination (but not to the extent they did with the other types of indirect meaning). As discussed in the previous section, we know that the use of a combination of strategies by low-scoring participants was often the result of them looking for additional cues to compensate for their inability to understand the target sentence. Whereas for high-scoring participants, these strategies were used as secondary/supportive sources.



Despite the mixed results from the low-scoring participants on these indirect refusals items, the fact that our high-scoring participants reached a 100% accuracy on these items during the interview and that they only needed to rely on one—sometimes two—strategies seem to indicate that this type of indirect meaning was relatively easy for them. In their study, Taguchi et al. (2016) also found that “comprehension was relatively straightforward with no need for using multiple strategies” (p. 691) for indirect refusals. Because indirect refusals were following a conventional pattern of refusal (that heavily relied on linguistic encoding), visual and auditory cues may not have been as salient and accessible to learners, which means that participants who were able to understand the target sentence-level meaning could process the conventional pattern without any issues. Additionally, the literal meaning of the target sentences for indirect refusals were not semantically opposed to the intended meaning (as it can be the case for meaning reversal irony), which means that the inferencing process for indirect refusals did not necessarily require the participants to detect an incongruity between what is said and what is implied. Participants who did not have access to the target sentence-level meaning, however, struggled with reaching the intended meaning due to the lack of paralinguistic cues (or the fact that they were not salient enough). Results from our high-scoring interview data support the idea that conventionality facilitates comprehension, as hypothesized by Taguchi et al. (2016—see also Dyzman and Ewert; 2010, Holtgraves, 1999, Taguchi, 2005, 2009a, 2011, etc.). Indeed, as other authors (Kuriscak, 2010; Rose & Kasper, 2001; Taguchi, 2009a; Taguchi & Yamaguchi, 2019) claimed before, once a certain threshold of L2 linguistic ability is reached, the conventionality effect can be detected and conventional indirect meaning becomes easier and faster to comprehend. And while we should use caution in generalizing these findings given the small sample, we also note that no significant difference in response times was found between indirect refusals and literal items during our in-class test, which could be explained by the fact that participants did not have to process multiple contextual sources to derive meaning (cf. Sperber & Wilson, 1995) for this type of indirect meaning category.

### 9.2.2. *Irony*

A total of 63 strategies was identified across 24 verbal reports focusing on irony, which presents the greatest number of strategies used among the three categories of indirect meaning. Accuracy rate for the low-scoring group during the interview was 58.33% for irony and the high-scoring group reached 91.67% accuracy.

Just like with indirect refusals, high-scoring participants used logical reasoning more than low-scoring participants to interpret ironic comments and low-scoring participants used keyword inferencing more (which we know can be attributed to the difference in proficiency level and ability to understand the target sentence). However additional interesting patterns of strategy selection emerged for this specific category of meaning. Overall, we note a greater reliance on paralinguistic cues (auditory and visual), background knowledge and setting/situation than for the other two types of indirect meaning (cf. [Table 23](#)). Both low-scoring and high-scoring participants also relied on a combination of strategies more often than for the indirect refusals (and to a lesser extent than for indirect opinions). Unlike indirect refusals, which are conventional and may be interpreted directly with the use of logical reasoning (as seen previously), irony presents a “great[er] distance between the two levels of meaning (sense and force) [which] requires an extensive amount of inferencing” (Taguchi et al., 2016, p. 692). We know that to detect irony, people must notice the incongruity between what is said and the situation or contextual cues (cf. [sub-section 2.6.2](#)) through irony markers (Attardo et al., 2013; Attardo & Poggi, 2003; D’Imperio et al., 2013; Gibbs, 2000; Haiman, 1998; Shively, 2013; Wallmark, 1971; Yus, 2000, 2017—cf. [section 2.6](#)). The fact that our participants relied on a combination of auditory, visual, situational cues, and background knowledge is indicative of that incongruity detection process, which our verbal reports clearly illustrate. We will now analyze some excerpts from our interviews to present some of the observed patterns and processes we just mentioned. As a reminder, the three examples of irony selected for the interviews were based on the following:

- Video Q21 (IRO 5) and video Q30 (IRO 14) were examples of a reversal of meaning: a conventional rhetorical device of saying the opposite of what is intended.
- Video Q24 (IRO 8) followed a more complex pattern of meaning replacement,<sup>67</sup> assimilated to an adynaton which we discussed in [Chapter 2](#).

---

<sup>67</sup> which Taguchi et al. considered conventional but which we argued did not allow for a very predictable pattern of interaction, thus making this type of ironic comments more difficult to interpret.

For convenience, we are including these three video scripts here:

***IRO 5 - Q21*** This is a conversation between two friends, Richard (male speaker) and Alicia (female speaker). They are talking about the trip Richard is planning to make.

***Dialogue:***

Richard: Alicia, est-ce que je t'ai dit que je vais partir en voyage le mois prochain?

Alicia: Non, je ne savais pas que tu allais partir en vacances! Où est-ce que tu vas aller?

Richard: Plusieurs villes en Europe: Amsterdam, Barcelone, Milan, Paris, Athènes, Berlin et Prague.

Alicia: Oh non, arrête, tu as la vie dure!

***IRO 14 - Q30*** This is a conversation between two classmates, Jean-Charles (male speaker) and Anabelle (female speaker). They are talking about how cold it is in class.

***Dialogue:***

Jean-Charles: Anabelle, c'est génial le printemps est arrivé! J'en avais vraiment besoin!

Anabelle: Oui, Jean-Charles. Mais je dois quand même mettre une veste en classe.

Jean Charles: Ah bon pourquoi? Il fait bon dans la salle de cours!

Anabelle: Ah oui bien sûr. On a l'impression d'être dans un sauna!

***IRO 8 - Q24*** This is a conversation between two friends, Paul (male speaker) and Iris (female speaker). They are talking about their college applications.

***Dialogue:***

Paul: Salut, Iris! Est-ce que tu as déjà envoyé des demandes d'inscription aux grandes écoles?

Iris Oui je les ai envoyées la semaine dernière! Et toi? Des nouvelles?

Paul: Non, mais je suis sûr que je vais être accepté partout. Je suis un excellent étudiant.

Iris: Oui, c'est ça, et moi je suis Harry Potter.

In IRO 5, Alicia's last comment is meant to be interpreted as the exact opposite of what she said: Richard's life is pretty good (based on the vacation he is planning and all the wonderful cities he is planning to visit). As explained by Taguchi et al. (2016), background knowledge was particularly important to understand the ironic comment in this dialogue because participants needed to understand that "these cities were all ideal places to visit" (p. 693) to detect the incongruity between Alicia's proposition and the situation. Additionally, paralinguistic cues should have been a salient source of access to irony comprehension: Alicia used a noticeable intonation when uttering the target sentence and, in the video, one can see her

make facial expressions (raise her eyebrows, widen and roll her eyes).

In IRO 14, Anabelle, expresses being cold in the classroom (and explained she has to wear a jacket in class, while on screen, the actress pulls her cardigan closer). As Jean-Charles comments that it feels quite good in the room, she sarcastically replies that it feels like a sauna. Here, we are again meant to understand the exact opposite: that she does not feel like it is hot like a sauna, but rather cold. Anabelle's body language was particularly important here as we see her grabbing onto to her sweater and clearly showing signs of being cold. Personal experience and background knowledge would indicate that when someone wears a sweater or jacket, they are usually cold. One would also need to be aware of the fact that saunas are actually very hot to understand the incongruity in Anabelle's sarcastic comment. In this video as well, the actress used a sarcastic tone and similar facial expressions as for the IRO 5 video.

IRO 8, is based on a rhetorical device called adynaton, which is essentially an extreme form of hyperbole using logical fallacy, in which "the exaggeration is taken to a ridiculous and literally impossible extreme, like "when pigs fly!" or "when Hell freezes over."<sup>68</sup> In IRO 8, when Paul brags about being an excellent student, Iris's comment that she is Harry Potter is meant to ironically provide a negative evaluative judgement of Paul. Indeed, just as Iris is obviously not Harry Potter, we are meant to infer that (she believes that) Paul is equally not an excellent student. Several authors (Alazzawie, 2016; Haiman, 1998) have studied this pattern where the sarcasm producer using the expression "And I am the Queen of England/Sheba" to express sarcasm (which follows a similar pattern as our "I am Harry Potter" example). While not written in a question form, this type of logical fallacy argument has also been compared to the "Pope Question."<sup>69</sup> Background knowledge would be needed to understand that Iris is not in fact Harry Potter, but some participants also expressed being familiar with this specific type of irony device (adynation) and understanding its mechanisms of production based on their experience in L1. Here again, paralinguistic cues were available (eye squinting, head shake, sarcastic tone, etc.). This particular irony item proved to be particularly difficult for participants during our in-class test: a special segment will be devoted to its analysis at the end of this sub-section.

[Excerpt 17](#) (which we already reviewed in [sub-section 8.4.4](#)) focuses on IRO 5 and is a great

---

<sup>68</sup> From <https://pan-gramma.com/blog/?p=16>

<sup>69</sup> Which Bouton (1994a) defines as a formula by which one answers a yes/no question with another rhetorical question to which the answer is obvious, and identical to the original question.

example of the use of multiple strategies (some based on primary sources, some on secondary sources) by a low-scoring participant, who happened to understand the target sentence, and used logical reasoning coupled with background knowledge to detect the incongruity (salient/leading sources—cf. Yus, 2000), which was then reinforced by visual and auditory cues (supportive sources—cf. Yus, 2000).

### Excerpt 17

*Video Q21 - IRO 5 - Participant 5 [low-scoring]*

**Researcher:** Alright, so I don't know if you are able to understand her last comment. But what are you, what are you thinking, out of the four options?

**Participant 5:** Um, so I know she said something along the lines of “*vous avez la vie dure*” which is you must have a hard life, but he just mentioned, going to all these different countries on vacation and it definitely seemed like sarcasm.

**Researcher:** Mhm.

**Participant 5:** So with that context. (4.0) Um, but that context, and with it, it just seems so sarcastic. So I'm really drawn to pick answer one. If we're going with literally it'd be two. But with the context, number two, wouldn't make any sense.

**Researcher:** Okay.

**Participant 5:** It's clearly not hard. He's about to tour freakin Europe, I wanna tour Europe.

**Researcher:** Yeah

**Participant 5:** I want to go to Europe.

**Researcher:** Yeah, that's exactly correct. Um were there other things, other cues that may have helped you decide that she was being sarcastic.

**Participant 5:** Definitely the body language.

**Researcher:** Body language and anything specific?

**Participant 5:** Like her eyes widen, like “*oh tu as la vie dure.*” Also it was also the intonation of her voice.

**Researcher:** Okay.

**Participant 5:** Which was obviously very kind of sarcastic.

**Researcher:** Mhm [nods]

**Participant 5:** Because the sarcasm from the intonation is what sealed the deal for me on number one.

While the low-scoring participant in [Excerpt 17](#) was able to understand the intended meaning easily based on the target sentence that he understood, low-scoring participants were not always successful in selecting the correct MCQ option, especially when they did not understand the target sentence (the two instances when they did, they reached the correct interpretation). Participants who did not have access to the target sentence-level meaning were only able to reach the correct answer half of the time. In these cases, they had to rely on other strategies (keyword, visual, auditory, setting, etc.) with a mixed success rate.

The 13 instances of visual cue strategies (six from the high-scoring and seven from the low-scoring group) were mostly successful. Participants reported the following facial expressions, gestures or body language, which they associated with sarcasm or irony: eyebrow raising, eye rolling, eye squinting, eye widening, head tilting, shrugging, snarky smile, disapproving head shake, as well as more general “body language.” These always combined with other strategies: when coupled with logical reasoning, they were used mostly as “supportive sources,” whereas participants had to rely on them as the “primary source” when they did not have access to the target sentence-level meaning (cf. Yus, 2000) for the interpretation of irony. Yus explains that (2000, p. 8):

It is commonly accepted that the speaker’s ostensive nonverbal behavior (smiles, frowning, winks, etc. belonging to the contextual information accessible to the addressee via perceptual mechanisms) can help the hearer to identify the intended ironic interpretation. In general, it is supposed that the speakers who want to facilitate the identification of an ironic message will foreground their nonverbal behavior as an explicit irony marker.

Auditory cues were also used at a fairly high frequency compared to the other meaning categories: we identified nine instances of use of this strategy for irony. As stated by Taguchi et al. (p. 692) we know that “irony is difficult to comprehend because the listener needs to attend to nonverbal cues (facial expressions, gestures, intonation) that make the utterance ironic,” and as such it seems that our participants followed that pattern of irony processing. Some of our participants reported that they could “**detect the sarcasm by the tone of [the characters’] voice**” (Participant 28 – high-scoring). The auditory cues were usually combined with other successful strategies and as such were rather effective at detecting irony, with the one exception that we already discussed in [sub-section 9.1.2.5, Excerpt 16](#), in which Participant 7 (low-scoring) misinterpreted the character’s attitude and intonation, and understood the target sentence in the complete opposite way.

Another major difference of strategy selection pattern for irony was the use background knowledge and setting/situation: we found nine instances of background knowledge and six instances of reliance on setting/situational cues, which was higher than for the other two categories of indirect meaning. This was particularly true for the comprehension of the two examples of irony that were based on meaning reversal (IRO 5 and IRO 14). These were used in combination with other strategies and were usually effective at interpreting the ironic comment. As discussed earlier, background knowledge was used for IRO 5 to recognize the

fact that the vacation Richard was planning was a great vacation, since he listed all the cities he would be visiting (for an example of this, see [Excerpt 17](#) above). Additionally, some participants used their life experience and knowledge of communicative interaction to draw parallels between the idiomatic expression “*tu as la vie dure*” and what English speakers could say in similar situations (as explained earlier with [Excerpt 3](#)<sup>70</sup>). As predicted, setting/situational cues and background strategies were also used for IRO 14 as participants often referred to seeing the character wearing a jacket (showing she was cold) in the video. The incongruity detected when the character mentions feeling like she is in sauna while also pulling her jacket closer and expressing the need to have to bring a jacket obviously relies on personal experience or common knowledge that one usually wears a jacket when they are cold, not hot.

We will now take a closer look at one specific irony item that seemed to have caused more difficulties for participants. Accuracy score for IRO 8 on the in-class test was the lowest of all items (across all four categories of meaning) with only 21.2% of participants selecting the correct answer.<sup>71</sup>

#### 9.2.2.1. Focus on IRO 8: the issue with Harry Potter

As mentioned before, the target sentence in IRO 8 is an adynation: a figure a speech that follows a specific pattern of communication (Alba Juez, 1998; Havertake, 1990) and as such participants in a conversation could be expected to interpret the adynation based on this prototypical pattern: someone says something that is deemed inaccurate or outrageous by another person who then replies with an even more outrageous comment (Yus, 2000). As mentioned before, background knowledge (or “common sense”—Yus, 2000, p. 5) is usually needed to recognize this “conventionalized ironic strategy” (which falls under “encyclopedic/factual information” in Yus, 2000), but Alba Juez also explains that this pattern follows “two successive turns organized into so-called adjacency pairs” (1998, p. 14). We would therefore expect the adjacency pair rule to be used as part of the inferencing process for this specific irony item (as we will see shortly with [Excerpt 18](#)).

---

<sup>70</sup> Q 21 – IRO 5 – Participant 3 : « Oh yeah « *tu as la vie dure* ». Like, it just looks like, you know, like the same way that people, English-speaking people do it like, “oh yeah, you have a hard life, don’t you”?”

<sup>71</sup> As a point of reference, overall accuracy rate for the Irony category was 65.66% during the in-class test.

During our interview, three of the high-scoring participants immediately selected the correct MCQ option using logical reasoning (they understood the target sentence) and a combination of other strategies (auditory, background, visual and/or adjacency pair rule). These participants were also familiar with the adnation mechanisms and were able to verbalize the process as shown in [Excerpt 18](#).

### Excerpt 18

*Video Q24 - IRO 8 - Participant 29 [high-scoring]*

**Researcher:** Alright, so what do you think, which number?

**Participant 29:** Uh, the... Iris does not think Paul is an excellent student. The third one.

**Researcher:** Okay, so tell me how you got to that answer.

**Participant 29:** Because he said that he can get into any school and she was like, "Yeah, and I'm Harry Potter" and she's clearly not Harry Potter.

**Researcher:** Yeah, so basically by making a statement... So she's making an affirmative statement saying that she's Harry Potter, which is obviously not true.

**Participant 29:** Yeah.

**Researcher:** And so she's implying that... what is she implying by doing that?

**Participant 29:** That he is not a good student. Because if he was a good student then she would be Harry Potter.

**Researcher:** Exactly, exactly. Um, any other cues that you noticed there?

**Participant 29:** She's kind of like smiling at him, she's kind of like snarky at the end.

**Researcher:** Yeah

**Participant 29:** Like she's shaking her head.

In [Excerpt 18](#) we can see Participant 29 first referencing the prototypical pattern of exchange: Paul mentioned that he is going to get into any school he wants [he claims to be an excellent student] and it is that statement that triggered Iris' sarcastic comment. We coded this as a strategy based on the adjacency pair rule. Participant 29 clearly understood the target sentence-level meaning and was able to accurately translate it and noticed the violation of the felicity condition (Havertake, 1990): Iris is "clearly not Harry Potter." When asked if he noticed other cues, Participant 29 mentioned attending to some visual cues but it seems these were supportive sources for him. A similar inferencing process was noted in the two other verbal reports in which the high-scoring participants reached the correct interpretation.

Participant 31 was the only high-scoring participant who was not able to select the correct answer for this item. She had also failed to reach the correct interpretation during the in-class test. During the interview, she admitted not being able to understand the target sentence, and



while she heard the male character, Paul, state that he was a really good student, she could not determine whether the female character, Iris, was being sarcastic or sincere in the last comment (target sentence). While she could not commit to selecting a MCQ option, she acknowledged the fact that the female character in many of the videos (even though the actress was supposed to play different characters) “would kind of give him [the male character] a hard time in a lot of videos so it kind of made me like suspicious that like she might be sarcastic” (Q24, IRO 8, Participant 31). This would be an example of a participant relying on their familiarity with the characters to try to interpret their comments. Nevertheless, it seems that the main reason why Participant 31 struggled with this video was because she could not access the literal meaning of the target utterance. As a matter of fact, during the second part of that verbal report (off-script, which was not coded for inferencing strategy—cf. [section 8.5](#)), we translated the target sentence for Participant 31, and she was then immediately able to reach the correct interpretation and explain the sarcastic process involved in the adynation.

We will now look at the low-scoring participants’ performance on this item. IRO 8 proved to be challenging for the low-scoring students: none of them were able to understand this video correctly. All four participants admitted that they did not know what the target sentence was saying. In fact, Participant 2’s justification for eliminating the MCQ option that referenced Harry Potter was that she did not hear Harry Potter mentioned in the video. We did not anticipate that participants may run into this comprehension issue: Indeed, this particular dialogue was based off of Taguchi et al.’s script (they also used the same Harry Potter reference at the end). The assumption was that this would be an easy reference for participants to access since they were all young adults and likely familiar with the fictional character. Our verbal reports revealed that, just like high-scoring Participant 31 (which we discussed above), low-scoring participants could not recognize the name Harry Potter due to the French pronunciation. Wanting to find out whether participants would be able to access the intended meaning once they had access to its literal meaning, we translated the target sentence into English. Like Participant 21, all but Participant 2 recognized the conventionalized pattern of the adynation and understood the ironic intent of the utterance.

In summary of this sub-section on irony, we offer some parallels between Taguchi et al.’s findings and ours. We know from our quantitative analysis (sub-sections [7.2.1](#) and [7.2.2](#)) that irony was challenging for L2 learners: accuracy was significantly lower than for other

categories of items and participants took longer to answer. Our results support other authors' (Bouton, 1992, 1994b; 1999; Kim, 2013, 2014, 2016; Shively et al., 2008; Taguchi et al. 2016) findings that irony presents a challenge for L2 learners. Slower comprehension speed can likely be explained by the need for participants to process multiple cues in order to detect and process the incongruity of the ironic utterance. As Taguchi et al. state, "participants' difficulty in comprehending irony is not surprising, considering that in irony the intended meaning is most remote from the propositional meaning (complete opposite of what is intended)" (2016, p. 692). It was obvious from verbal reports that the primary criteria for understanding ironic utterances was the ability to access the target sentence-level meaning, but understanding the literal meaning of the target sentence was not enough—they needed to look for clues within the context to detect, process and interpret the ironic comments:

Without processing these multiple cues, i.e., linguistic, nonverbal, and contextual, learners might wrongly jump to the literal meaning. [...] To comprehend irony, learners attended to a variety of cues that were salient to them, including paralinguistic cues, contextual cues, and visual information, to arrive at sarcastic meaning.

Taguchi et al., 2016, p. 693

The presence of paralinguistic cues in our irony videos is congruent with research that shows that visual and auditory cues are often used as irony markers by speakers (Attardo, 2000b, 2002; Attardo et al., 2003; Giora & Fein, 1999; Yus, 2000, 2017—cf. sub-section 2.6.4 for all other references). These cues were used primarily as supportive source of ironic interpretation by our participants. Our interview data does not support Taguchi et al.'s claim that "[a]lthough some learners were successful in comprehending irony, it was this multimodal processing that added to the comprehension difficulty, resulting in the significantly lower accuracy score of irony compared with other item types" (2016, p. 693). In our opinion, it was the inability of the participants to access the literal meaning of the target utterance that prevented them from accessing the intended meaning. Salient audiovisual cues were actually used successfully by participants – particularly by the high scoring participants – as a way to confirm the incongruity they detected based on the context (using background knowledge, previous utterances, and literal meaning as leading sources). These findings are consistent with what Shively et al. (2008) observed in their study on irony interpretation: audiovisual cues were most beneficial to their advanced proficiency learners. In the current study, when participants could not understand the target utterance, these auditory and visual cues were sometimes used as compensatory tools to infer meaning (with mixed success). In this regard, L2 listening

comprehension appears to have been the determining factor (as seen in Taguchi, 2008a, 2009a) which was likely linked to our participants' proficiency level.

### 9.2.3. *Indirect opinions*

A total of 48 strategies were identified in the 24 indirect opinion verbal reports. Just like with indirect refusals (but not to the same extent), we observed several instances of uses of a single strategy for indirect opinions (which Taguchi et al. also observed). However, this phenomenon occurred in both low-scoring and high-scoring participants (4 instances for each group). We should note here that while indirect opinions had the highest accuracy score (78.28%) among the three categories of indirect meaning on the in-class test, the three items that were selected for the interviews had a much lower accuracy rate on the test (53.55% across all participants). This was also reflected in our verbal reports: our participants were only able to select the correct MCQ option for these three items 54.16% of the time during the interviews (accuracy for low-scorers was 33.33% and for high-scorers 75%). We will therefore use caution when we interpret the results from our interview data as they may not be representative of the difficulties encountered by participants with the whole set of indirect opinion items, or their inferencing processes on the rest of the items.

In [Chapter 4](#) (sections [4.1.2](#) and [4.5](#)), we discussed the fact that some of the indirect opinion items from Taguchi et al. seemed to adopt a sarcastic mode of communication. The target sentences in the three items selected for the interview all represent an indirect negative evaluative judgement (cf. [Appendix C](#)). In OPI 2, Anna indirectly criticizes a bride's gown, in OPI 7 Sophia indirectly expresses dissatisfaction with a massage she just received, and in OPI 8 Florence essentially implies the movie she watched was boring. Each time, the dialogues observed a similar pattern of turn-taking: the male character asks for the female character's opinion regarding the gown, massage and movie, which she indirectly provides. While Taguchi et al. (2016) claim that indirect opinions are not conventional, we know that requests for evaluative judgement from the male character set some expectations for the next turn, and as such may follow a predictable pattern of interaction (cf. adjacency pair, [section 2.7](#)).

Just like for the other two categories of indirect meaning, the low-scoring group used keyword inferencing more than the high-scoring group and the high-scoring group used logical reasoning more than the low-scoring group (cf. [Table 24](#)). However, keyword inferencing was actually

the most frequently used strategy within the category of indirect opinions (with 15 occurrences, which represented 31.3% of all strategies used for indirect opinions) and compared to the other two categories of indirect meaning. Indeed, even high-scoring participants had to rely on this strategy (five times) when they could not understand the target sentence. This was particularly true for OPI 2 and OPI 7. We have already discussed the fact that keyword inferencing was not a very effective strategy in [sub-section 9.1.2.1](#) and provided examples of both low-scoring and high-scoring participants' use of the strategy ([Excerpt 9](#), in [sub-section 9.1.2.1](#) actually illustrates a high-scoring participant's failure to understand an indirect opinion due to using keyword inferencing as a single strategy). The three instances of a single use of keyword inferencing by high-scoring participants were responsible for all three incorrect answers for this participant group in the indirect opinion category (and for the overall low accuracy score in this category). Reliance on this keyword strategy was therefore likely the reason why accuracy rate was so low and can be explain by the difficulties for participants to access the literal (linguistically encoded) meaning of the target utterance (cf. [sub-section 9.1.1](#)).

Another notable difference for this category of indirect meaning was the use of visual cues: we identified 13 instances of use of this strategy (same number as for irony) which represented 27.1% of all strategies used for indirect opinion comprehension during our interviews (see [Table 23](#)). In [sub-section 9.1.2.3](#), we provided an example of a low-scoring participant who was able to select the correct MCQ option for an indirect opinion item solely based on visual cues ([Excerpt 12](#)). However, the visual cues were often used in combination with other strategies (in a successful manner with logical reasoning—most often by high-scoring participants—and with keyword inferencing by low-scoring participants). Taguchi et al. also found that their participants “used visual cues much more often when comprehending indirect opinions.” (p.691). They believed the greater use of visual cues for indirect opinions was responsible for the faster comprehension speed on their test (p.691):

This tendency could explain why learners' comprehension was faster for indirect opinions than for refusals. Different from logical reasoning that requires processing of individual utterances, use of visual cues is more economical because visual information directly maps onto meaning. As a result, listeners can bypass the bottom-up processing of an utterance, which results in shorter response times.

This claim was not supported by our results: even though our participants (like Taguchi et al.'s) used more visual cues and fewer logical reasoning strategies to comprehend indirect opinions

than indirect refusals, unlike Taguchi et al., comprehension speed for indirect opinions was slower than for indirect refusals in our data.

The analysis of our verbal reports did not provide any explanation for the relatively high accuracy rate of the indirect opinion category on the test. However, we should remember that the three indirect opinion items we selected for the interview were not necessarily representative of the difficulties for this category level (accuracy on these three items was much lower during the interview) and as a result, the patterns of strategies used during the interviews may not be typical of the inferencing processes used by L2 learners for this category of meaning in general.

The fact that interview participants struggled with the indirect opinions item during the interview and were not as successful in selecting the correct MCQ option (as compared to the other two categories of meaning) could lend some support to the Taguchi et al.'s claim that conventionality impacts L2 comprehension of indirect meaning, as according to them, nonconventional implicatures like indirect opinions would be harder to comprehend. However, in light of what we just said regarding our interview item sampling, we will refrain from making this type of generalization. Given the low accuracy and high number of keyword inferencing strategies, we argue that our L2 learners' comprehension difficulties may have stemmed from their inability to access the literal meaning of the target utterances more than from the potential lack of conventionality of this category of indirect meaning.

# **GENERAL CONCLUSION**

## Summary of findings and importance of replication

Our study investigated the comprehension of indirect meaning by L2 French learners, replicating the methodology used by Taguchi et al. (2016) with L2 Spanish learners. Although the purpose of our close replication was to “test generalizability of the earlier findings under different conditions” (Marsden et al., 2018a, p. 322)—validating Taguchi et al.’s findings with a different target language—we introduced a few changes to “increase the transparency and robustness of the analysis” (Porte & McManus, 2019, p. 121). To answer our first research question (To what extent are L2 French learners able to understand indirect meaning?), we opted to include the fillers in our quantitative analyses so that we could have a point of comparison between indirect and literal meaning comprehension. We also reprocessed our native speaker dataset to only include items that were used in the L2 learner test so that we could better compare comprehension of L2 learners to that of our native speakers’. As a result, we were able to respond to our first research question from two angles: the extent to which L2 learners are able to understand indirect meaning as compared to native speakers, and the extent to which they are able to understand indirect meaning as opposed to literal meaning. Our data show that our L2 French learners did not perform as well as our native speakers (the former only reached 72.64% accuracy overall on the three indirect meaning categories, whereas the latter were accurate 99.52% of the time on these same items). Unlike our native speakers (who comprehended all categories of meaning at a very high and consistent level—i.e., no statistical differences were found in native speaker accuracy and speed on the retained items), our L2 learners showed some differences in comprehension accuracy and speed between some of the indirect meaning categories (specifically irony) and the literal meaning items (the latter received the highest accuracy score of 83.33%). However, not all indirect meaning categories were found to be significantly more challenging than literal items. This leads us to our second research question: Do different types of indirect meaning impact L2 French learners’ comprehension accuracy and speed? As stated above, our statistical tests revealed that irony was more difficult: we found significant differences between irony and all other categories of meaning (indirect refusals, indirect opinions and literal items) in both accuracy and speed. In other words, our L2 learners did not understand irony items as well as the other categories and were slower to process those examples. The only other statistical difference we found was between indirect refusals and literal meaning: our L2 learners were less accurate in interpreting indirect refusals than literal meaning, but the effect size for this comparison was small.

The analysis of our verbal interviews allowed us to answer our third research question: What inferential strategies do L2 French learners use and how successful are they? The identification of the various inferencing strategies in our data revealed some interesting patterns of strategy selection by L2 learners, which highlights a complex multi-dimensional comprehension process. Our data supports a strong reliance on contextual cues, consistent with Sperber and Wilson's (1995) Relevance Theory, which Taguchi et al. also underline in their paper (2016, p. 693):

The variety of strategies that participants mentioned using in the verbal protocols supports relevance theory (Sperber & Wilson, 1995), which claims that utterance interpretation is neither a mere decoding of linguistic input nor retrieval of logic. Various kinds of information in context, both linguistic and nonlinguistic, are placed on the same level to assist comprehension. The video-based task used in this study gave learners a context in which they could use prosodic, discourse, and nonverbal cues to assist their interpretation of ironic utterances.

However, these contextual cues were not always used in the same way by the high-scoring and low-scoring participants—a difference that was brought to light by our group comparison analysis. Indeed, our high-scoring participants were able to access the literal meaning more easily than our low-scoring participants (they demonstrated a higher level of listening/linguistic proficiency), and they use contextual cues to draw inferences about the intended meaning. When used as a supporting source for inferencing (combined with access to the literal meaning), these strategies were extremely effective in reaching the correct interpretation. On the other hand, our low-scoring participants struggled to access the literal meaning of the target utterance. As a result, the strategies they used were more guessing/compensatory strategies than actual inferencing strategies (in the sense of deriving meaning from the linguistic input to arrive to the intended meaning, the way native or high-proficiency speakers would). This is in line with previous research (cf. Rost & Ross, 1991; Tsui & Fullilove, 1998; Zoghiami, 2016; Zoghiami & Hilton, 2021) outlined by Field (2008, p. 132) as follows:

The truth is that both unskilled and skilled listeners make use of context, but that they do so for different purposes.

- Skilled listeners and readers (whether in L1 or L2) make use of context to enrich their understanding of the message. Less skilled listeners are not always able to achieve this wider understanding because their attention is so heavily focused upon details of the signal.



- Less skilled listeners and readers (whether in L1 or L2) make greater use of context and co-text to compensate for parts of the message that they have not understood. In second language listening, the failure might be due to problems of decoding, problems of word and grammar knowledge or problems in recognising the relationships that link ideas.

While our study did not focus on L2 listening skills only (since our multimodal instrument included a visual dimension to the comprehension process), our findings seem to support the Linguistic Threshold Hypothesis that presupposes a certain level of L2 linguistic ability must be reached to achieve successful listening comprehension (Vandergrift, 2006). Our less proficient L2 learners had to compensate for their poor linguistic ability and rely on other cues. This would explain why our low-scorers used a greater number of strategies overall than our high-scorers (cf. interactive-compensatory mechanism—Field, 1998)—a result that was opposite to what Taguchi et al. had found. These compensatory strategies were not always successful,<sup>72</sup> and results from our low-scorer data support Zoghiami (2016) and Zoghiami and Hilton’s (2021) findings with regard to listening comprehension (i.e., that lower-proficiency learners do use top-down strategies, but that a certain level of linguistic knowledge is necessary for successful oral<sup>73</sup> comprehension). Yet, the fact that our low-scorers understood so little (at least linguistically) but still managed to select the correct MCQ options 47.22% of the time does highlight the (relative) efficacy of top-down strategies in L2 comprehension.

The difference between our three indirect meaning types (and its potential impact on comprehension) can also be seen in the frequency and variety of strategies used by our two groups of learners for certain categories of items. Focusing on the performance of our high-scoring participants during the interview, we noticed that they not only understood indirect refusals better but they also used fewer strategies, which would indicate that overall, this type of indirect meaning (categorized as conventional by Taguchi et al.) is easier for high proficiency learners to process. On the other hand, indirect refusals seemed surprisingly challenging for low-proficiency learners during the interview. We suggest that this might be due to the fact that audiovisual cues may not have been as present or salient enough for this type of indirect meaning. As a result, our lower proficiency learners, who were unable to understand these

---

<sup>72</sup> cf. Rost (1990)’s view on editing strategies and the fact that vocal and visual cues when used alone (without access to propositional meaning) “provide the basis for weak inferences about speaker meaning” (p. 79).

<sup>73</sup> Zoghiami, 2016; Zoghiami & Hilton, 2021 only focused on listening comprehension without the visual element.

items, could not rely on other cues to interpret them. Again, this is consistent with research on listening comprehension (cf. Zoghiami, 2016; Zoghiami & Hilton, 2021).

With regard to conventionality, we argued that a possible reason why a conventionality effect was not detected in our analyses may lie in the nature of our multimedia tasks. Taguchi et al. suggested that indirect opinions (nonconventional) comprehension was facilitated by the access to visual cues present in the videos (which resulted in higher levels of accuracy compared to previous studies). Our data support this explanation: visual cue strategies were used for indirect opinions more than for the other two types of indirect meaning. This would indicate that the conventionality effect may be task-dependent and could be offset by the presence of paralinguistic cues. Whereas learners may be able to understand conventional indirect meaning like indirect refusals solely on the basis of the linguistically encoded meaning (without having to rely on visual cues), the presence of paralinguistic cues may allow them to reach the intended meaning of less conventional indirect meaning (sometimes even when they do not understand the literal meaning of these items).

As previously stated, one of the reasons we chose to replicate Taguchi et al.'s study was because they reported an unexpected finding: whereas previous research on indirect meaning comprehension showed that conventionality facilitated comprehension, Taguchi et al. did not find significant differences between indirect refusals (conventional indirect meaning) and indirect opinions (nonconventional indirect meaning). As McManus (2021) stated, "replication is the principal way to move us towards a more nuanced, finer-grained understanding about the nature and a validity of a specific study's findings" (p. 2). Our results allowed us to increase the confirmatory power of the original study and reinforce the validity of Taguchi et al.'s findings with regard to the conventionality effect (i.e., our L2 learners did not understand indirect refusals significantly better than indirect opinions either). Our study also corroborated previous research findings (including Taguchi et al.'s) with regard to irony comprehension. Indeed, this type of indirect meaning was the most challenging for our L2 learners who were less accurate and took longer to process irony items. In addition to being able to verify some of Taguchi et al.'s findings, we also contributed to advancing our understanding of L2 comprehension by expanding the scope of our analyses by adopting field-appropriate effect size recommendations, including literal items, and looking at the interactions of proficiency, indirect meaning type and task in our interview data. The intersections of these three factors were not necessarily (or sufficiently) examined by Taguchi et al. Ultimately, results from our interview

data suggest that difficulties with indirect meaning comprehension may be less dependent on the indirectness of meaning than on more global oral comprehension difficulties.

## Pedagogical implications

Given the relative difficulty L2 learners experienced with irony, this type of indirect meaning may require explicit classroom instruction. To that effect, Taguchi et al. also recommend some pedagogical activities to improve irony detection and comprehension. The following are based on their recommendations as well as our own suggestions:<sup>74</sup>

- Present examples of irony, identify and explain specific patterns when present (such as meaning reversal, adynation, Pope question, etc.)
- Increase awareness of paralinguistic cues that often accompany irony (facial expressions, sarcastic tone, etc.) through the use of video-based tasks (identify cues, catalog them, explain possible in-context interpretation, etc.).
- Encourage learners to use both top-down (paralinguistic and contextual cues) and bottom-up (linguistic cues from literal meaning) strategies (i.e., pay attention to paralinguistic cues and not solely rely on keyword-inferencing strategies).
- Include cross-linguistic comparisons: ask learners how irony is expressed in their own language and culture (both verbally and non-verbally), analyze those patterns and compare with the target language/culture.

More generally speaking, instructors should also discuss the way indirectness is used to convey certain messages based on contextual and cultural appropriateness (situation, interlocutors' status, politeness, etc.) as well as its communicative impact (how indirectness may reflect the speaker's attitudes and emotions, and how it may affect the hearer and conversational exchange).

Since the main factor in our L2 learners' inability to understand indirect meaning was their lack of basic linguistic understanding, we must emphasize the need to develop overall proficiency (phonetic, lexical, grammatical, parsing, etc.) so that L2 learners are able to use both bottom-up and top-down processes (cf. Aponte-de-Hanna, 2012; Graham, 2017; Zoghiami, 2016; Zoghiami & Hilton, 2021). Increasing learners' linguistic and listening proficiency should decrease the need for compensatory strategies (cf. Rost & Ross's "proficiency position," 1991,

---

<sup>74</sup> Also similar to Kim (2013, 2014, 2016)'s recommendations for teaching conceptual understanding of sarcasm.

p. 263). However, our qualitative data also show that whenever proficiency is low, L2 learners can rely on contextual cues to interpret indirect meaning. This supports some authors' recommendation to teach appropriate and effective top-down strategies based on the type of tasks L2 learners are trying to accomplish (Graham, 2017; Rost & Ross, 1991; Vandergrift, 2006). In the case of indirect meaning comprehension, using multimedia (audio-visual) materials and tasks presents a clear advantage: since it "more closely reflects naturalistic processing because it makes a greater number of cues available to listeners and thus presents a more realistic approach to understanding L2 comprehension of indirectness" (Taguchi et al., 2016, p. 682). The addition of the visual elements (as opposed to just listening or even written tasks) means that L2 learners are able to rely on those paralinguistic cues that often accompany or convey indirect meaning, and instructors can use the videos to develop L2 learners' inferencing or compensatory skills.

### **Limitations and suggestions for future research**

McManus (2021) explains that a replication "allows us to understand how a study's research data were collected, measured, and analyzed, as well as the extent to which unexpected and/or unanticipated factors potentially shaped the conclusions" (p. 3). While undertaking this close replication of Taguchi et al.'s study, we had to carefully examine how their study was carried out, a process that led us to adopt a more critical (but informed) posture with respect to the study design and methodology. Some of these problematic aspects have been discussed in previous chapters, but we would like to highlight some of our study's limitations and offer some potential solutions when possible.

First, we recognize the fact that a multiple-choice questionnaire does not accurately assess participants' ability to understand indirect meaning. Not only is a MCQ not representative of naturalistic interactions (hearers are not usually provided with four options to choose from in normal conversations), but the format of our data collection instrument provided some elements of interpretation to the participants (as evidenced by the use of the process of elimination strategies). Even when our L2 learners were able to select the correct MCQ option, it does not mean they would have reached the correct interpretation if they had been asked to verbalize it without the MCQ. One way to remedy this issue would be to conduct on-line<sup>75</sup> verbal interviews

---

<sup>75</sup> On-line in the sense of "measurements taken concurrent to task performance," Veenman (2011), p. 206.  
268

(e.g., think-aloud protocols—cf. Zoghiami, 2016<sup>76</sup>), asking participants to explain how they interpret indirect utterances after each dialogue (using open-ended questions such as “what do you think the character means?,” “how do you interpret this last comment?,” etc.).

Another limitation concerns the response times. While we tried to control for length of MCQ options, we cannot account for differences in participants’ reading speed, or the fact that participants may not read all MCQ options (e.g., if they think the first option is correct, they may select it without reading the others, resulting in a faster reaction time). Even though we randomized the MCQ options for each participant, these observations lead us to suggest that response times are not always an indication of comprehension speed or a sign of comprehension difficulty.

As previously mentioned, when it comes to our verbal retrospective interviews, we should remember that qualitative data is often “open to some degree of subjective interpretation” (Ross, 1997, p. 236). Additionally, verbal introspections also have limits insofar that they rely on participants’ ability to verbalize their thought process. This means that interview data may only provide an incomplete picture of the participants’ comprehension processes. We suspect that our participants attended to additional cues, even possibly without realizing it, but never mentioned it. This is likely true for the strategy involving adjacency pair rule: since our participants knew that we were focusing on the last utterance in the dialogue, they rarely referenced the previous turn, even though it may have been part of their inferencing process.

While we did include the fillers (literal meaning items) in our quantitative analyses (which Taguchi et al. did not do), we believe that further investigation is needed to examine the differences in pragmatic comprehension between direct and indirect meaning. In the context of our replication, we did not analyze our retrospective interview data on the literal items (because we were focusing on identifying inferencing strategies, i.e., strategies used to infer indirect meaning). However, an analysis of the L2 learners’ comprehension processes for the literal items would offer a good point of comparison between strategies used for literal meaning and strategies used for indirect meaning (though we would have to use different coding categories at that point) and may produce more evidence for our claim that our L2 learners’ difficulties had more to do with general listening comprehension than with indirect meaning. Additionally,

---

<sup>76</sup> Zoghiami used a series of recordings that were paused at specific point, at which point participants were asked to verbalize their efforts to construct meaning (2016, p. 4)

given that irony was more difficult, we suggest that a more in-depth study of L2 French learners' comprehension of irony may give us a more accurate idea of where learners struggle with this category of indirect meaning. To that effect, researchers should develop tasks that include similar dialogues, in which the target utterances may have either an ironic interpretation or a literal meaning (in the same way Bromberek-Dyzman & Ewert, 2010, and Bromberek-Dyzman & Rataj, 2016, conducted their experiments with Polish EFL students). Participants' proficiency should also be assessed and factored into the study, and we recommend assessing learners' comprehension of irony in both the target language (L2 French) and their native language. Such design would allow researchers to clearly determine whether L2 comprehension of irony is challenging simply due to linguistic input (impact of proficiency), because of the indirect nature of irony (greater distance between what is said and what is meant) or whether individual differences may also account for variations in comprehension and processing.

Lastly, we should recall that our close replication also led us to question the categorization of our indirect meaning items according to their conventionality. In addition to challenging Taguchi et al.'s classification, we should point out that both our project and theirs stemmed from the assumption that the L2 and L1 under study shared similar conventional patterns of conversation: however, we cannot generalize L2 indirect comprehension without taking the cultural aspect into account. Future research should therefore aim to expand the current study to other languages, especially those that may present different conventional patterns of indirect expression.

# REFERENCES

- Abbuhl, R. (2012). Why, when, and how to Replicate Research. In A. Mackey and S. M. Gass (eds.) *Research Methods in Second Language Acquisition: A Practical Guide, First Edition*, (pp. 296–312). Wiley-Blackwell.
- Ackerman, B. P. (1983). Form and Function in Children's Understanding of Ironic Utterances. *Journal of Experimental Child Psychology*, 35(3), 487–508.
- Agar, M. H. (2002). *Language shock: Understanding the culture of conversation*. Perennial.
- Akimoto, Y., Sugiura, M., Yomogida, Y., Miyauchi, C. M., Miyazawa, S., & Kawashima, R. (2014). Irony comprehension: social conceptual knowledge and emotional response. *Human brain mapping*, 35(4), 1167–1178.
- Alazzawie, A. (2016). A Pragmatic Study of Yaʿni, an Iraqi Arabic DM, and English equivalents, *European Journal of Language Studies*, 3( 2), 57–74.
- Alba Juez, L. (1998). On the conventionalization of verbal irony in English: A corpus-based study. In I. Vasquez Orta & I. Guillen Galve (Eds.), *Perspectivas pragmaticas en linguistica aplicada* (pp. 11–18). Anubar.
- Al-Duleimi, H. Y., Rashid, S. Md., & Abdullah, A. N. (2016). *A critical review of prominent theories of politeness*. Australian International Academic Centre.
- Aponte-de-Hanna, C. (2012). Listening Strategies in the L2 Classroom: More Practice, Less Testing. *College Quarterly*, 15(1), 1–9.
- Armstrong, N. (2002). Variable deletion of French ne. A cross-stylistic perspective, *Language Sciences*, 24, 153–173.
- Ashby, W. (1981). The Loss of the Negative Particle ne in French: A Syntactic Change in Progress. *Language*, 57(3), 674-687.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Chapter: Human memory: A proposed system and its control processes. In Spence, K. W., & Spence, J. T. (Eds.), *The psychology of learning and motivation* (Volume 2, pp. 89–195). Academic Press.
- Atkinson, R. C., & Shiffrin, R. M. (1971). *The control processes of short-term memory*. Institute for Mathematical Studies in the Social Sciences, Stanford University.
- Atlas, J. D., & Levinson, S. C. (1981). It-Clefts, Informativeness, and Logical Form: Radical Pragmatics (Revised Standard Version). In P. Cole (Ed.), *Radical Pragmatics* (pp. 1–61). Academic Press.
- Attardo, S. (2017a). The general theory of verbal humor. In S. Attardo (Ed.), *The Routledge handbook of language and humor* (pp. 126–142). Routledge.

- Attardo, S. (2017b). Humor and pragmatics. In S. Attardo (Ed.), *The Routledge handbook of language and humor* (pp. 174–188). Routledge.
- Attardo, S. (2000a). Irony as relevant inappropriateness. *Journal of Pragmatics*, 32(6), 793–826.
- Attardo, S. (2000b). Irony markers and functions: Towards a goal-oriented theory of irony and its processing. *RASK: International Journal of Language and Communication* 12, 3–20.
- Attardo, S. (2002). Humor and irony in interaction: From mode adoption to failure of detection. In L. Anolli, R. Ciceri, & G. Riva (Eds.), *Say not to say: New perspectives on miscommunication* (pp. 159–179). IOS Press.
- Attardo, S., Pickering, L., & Baker, A. (2013). Prosodic and multimodal markers of humor in conversation. In S. Attardo, M. M. Wagner, & E. Urios-Aparisi (Eds.), *Prosody and Humor* (pp.37–60). John Benjamins Publishing.
- Attardo, S., & Poggi, I. (2003). Multimodal markers of irony and sarcasm. *Humor* 16(2), 243–260.
- Attardo, S., & Raskin, V. (2017). Linguistics and humor theory. In S. Attardo (Ed.), *The Routledge handbook of language and humor* (pp. 49–63). Routledge.
- Austin, J. L. (1962). *How to Do Things with Words*. Harvard University Press.
- Barbe, K. (1989). Irony in spontaneous conversation. In R. M. Brend & D. G. Cockwood (Eds.), *The 15<sup>th</sup> Lacus Forum 1988* (pp. 271–280). Linguistic Association of Canada and the United States.
- Barbe, K. (1993). "Isn't it ironic that...": Explicit irony markers. *Journal of Pragmatics*, 20(6), 579–590.
- Barbe, K. (1995). *Irony in context*. John Benjamins Publishing Company.
- Bardovi-Harlig, K. (1999). Exploring the interlanguage of interlanguage pragmatics: a research agenda for acquisitional pragmatics. *Language Learning*, 49(4), 677–713.
- Bardovi-Harlig, K. (2009). Conventional expressions as a pragmalinguistic resource: recognition and production of conventional expressions in L2 Pragmatics. *Language Learning*, 59(4), 755–795.
- Bardovi-Harlig, K. (2012). Formulas, routines, and conventional expressions in pragmatics research. *Annual Review of Applied Linguistics*, 32, 206–227.
- Bardovi-Harlog, K. (2013). Developing L2 Pragmatics. *Language Learning*, 63(Suppl. 1), 68–86.
- Bardovi-Harlig, K. (2014). Documenting interlanguage development. In Z. Han & E. Tarone (Eds.), *Interlanguage: Forty Years Later (Language Learning & Language Teaching)* (pp. 127–146). John Benjamins Publishing Company.



- Bardovi-Harlig, K., & Bastos, M.-T. (2011). Proficiency, length of stay, and intensity of interaction, and the acquisition of conventional expressions in L2 Pragmatics. *Intercultural Pragmatics*, 8(3), 347–384.
- Bardovi-Harlig, K., & Dörnyei, Z. (1998). Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning. *TESOL Quarterly*, 32(2), 233–262.
- Bardovi-Harlig, K., & Vellenga, H. E. (2012). The effect of instruction on conventional expressions in L2 pragmatics. *System*, 40, 77–89.
- Bell, N., & Attardo, S. (2010). Failed humor: Issues in non-native speakers' appreciation and understanding of humor. *Intercultural Pragmatics*, 7(2), 423–447.
- Berrendonner, A. (1981). *Eléments de pragmatique linguistique*. Les Editions de Minuit.
- Bertrand, R., & Priego-Valverde, B. (2013). Does prosody play a specific role in conversational humor? In S. Attardo, M. M. Wagner, & E. Urios-Aparisi (Eds.), *Prosody and Humor* (pp. 141–165). John Benjamins Publishing.
- Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, 19(4), 207–217.
- Blum-Kulka, S., House, J., & Kasper, G. (1989). *Cross-cultural pragmatics: Requests and apologies*. Ablex.
- Bossut, H. (2011). *Evaluation de la cognition sociale: normalisation d'un test de compréhension des sarcasmes auprès d'une population de 50 à 65 ans*. [Ressource électronique] <http://pepite-depot.univ-lille2.fr/nuxeo/site/esupversions/a8072177-e486-4535-a4cc-4157bfd6e3fd>.
- Bouton, L. F. (1988). A cross-cultural study of ability to interpret implicatures in English. *World Englishes*, 17, 183–196.
- Bouton, L. F. (1990). The effective use of implicature in English: Why and how it should be taught in the ESL classroom. In L. F. Bouton & Y. Kachru (Eds.), *Pragmatics and language learning*, (Vol. 1, pp. 43-51). University of Illinois, Urbana-Champaign: Division of English as an International Language.
- Bouton, L. F. (1992). The interpretation of implicature in English by NNS: Does it come automatically without being explicitly taught? In L. F. Bouton (Ed.), *Pragmatics and language learning* (Monograph Series Vol. 3, pp. 64–77). Urbana-Champaign, IL: Division of English as an International Language, University of Illinois, Urbana-Champaign.
- Bouton, L. F. (1994a). Conversational implicature in second language: Learned slowly when not deliberately taught. *Journal of Pragmatics*, 22, 157–167.
- Bouton, L. F. (1994b). Can NNS skill in interpreting implicature in American English be improved through explicit instruction?: A pilot study. In L. F. Bouton & Y. Kachuru (Eds.), *Pragmatics and language learning* (Monograph Series Vol. 5, pp. 88–108). Urbana-

Champaign, IL: Division of English as an International Language, University of Illinois, Urbana-Champaign.

Bouton, L. F. (1999). Developing nonnative speaker skills in interpreting conversational implicatures in English. In E. Hinkel (Ed.), *Culture in second language teaching and learning* (pp. 47–70). Cambridge University Press.

Bromberek-Dyzman, K. (2015). Irony Processing in L1 and L2: Same or Different? In R. Heredia & A. Cieślicka (Eds.), *Bilingual Figurative Language Processing* (pp. 268–298). Cambridge University Press.

Bromberek-Dyzman, K., & Ewert, A. (2010). Figurative competence is better developed in L1 than in L2, or is it? In M. Putz & L. Sicola (Eds.), *Cognitive Processing in Second Language Acquisition: Inside the learner's mind* (pp. 317–334). John Benjamins Publishing.

Bromberek-Dyzman, K., & Rataj, K. (2016). Irony comprehension in the nonnative language comes at a cost. *Psychology of Language and Communication*, 20(3), 336–353.

Brône, G. (2008). Hyper- and misunderstanding in interactional humor. *Journal of Pragmatics*, 40, 2027–2061.

Brown, P., & Levinson, S. (1978). Universals in Language Usage: Politeness Phenomena. In E. Goody (Ed.), *Questions and Politeness: Strategies in Social Interaction* (pp. 56–310). Cambridge University Press.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.

Brownell, H.H., Jacobs, J.R., Gardner, H., & Gianoulis, D. (1990). Conditions for sarcasm. Unpublished paper from the *Boston University Aphasia Research Center and Harvard Project Zero*.

Bryant, G. A. (2013). Verbal irony in the wild. In S. Attardo, M. M. Wagner, & E. Urios-Aparisi (Eds.), *Prosody and Humor* (pp. 103–120). John Benjamins Publishing.

Bryant, G. A., & Fox Tree, J. E. (2002). Recognizing Verbal Irony in Spontaneous Speech. *Metaphor and Symbol*, 17(2), 99–117.

Burgers, C., van Mulken, M., & Schellens, P. J. (2011). Finding irony: An introduction of the Verbal Irony Procedure (VIP). *Metaphor and symbol*, 26, 186–205.

Carston, R. (1988). Implicature, explicature, and truth-theoretic semantics. In R. M. Kempson (Ed.), *Mental representations: The interface between language and reality*, (pp. 155–181). Cambridge University Press.

Chamot, A. U. (2005). Language learning strategy instruction: Current issues and research. *Annual Review of Applied Linguistics*, 25, 98–111.

Channon, S., Rule, A., Maudgil, D., Martinos, M., Pellijeff, A., Frankl, J., Drury, H., & Shieff, C. (2007). Interpretation of mentalistic actions and sarcastic remarks : Effect of frontal and posterior lesions on mentalising. *Neuropsychologia*, 45, 1725–1734.

Charaudeau, P. (2006). Des catégories pour l'humour ? *Questions de communication*, 10, 19–41.

Charaudeau, P. (2011). Des catégories pour l'humour. Précisions, rectifications, compléments. In D. Vivero Ma (Ed.), *Humour et crises sociales. Regards croisés France-Espagne*, (pp.9–43). L'Harmattan.

Chen, H., J. (1996). *Cross-cultural comparison of English and Chinese metapragmatics in refusal*. Doctoral Dissertation. Indiana University  
<http://files.eric.ed.gov/fulltext/ED408860.pdf>.

Chen, R. (1990). Towards a pragmatic account of irony. *Language and Style*, 23(1), 29–37.

Chen, X., Ye, L., & Zhang, Y. (1995). Refusing in Chinese. In G. Kasper (Ed.), *Pragmatics of Chinese as native and target language. Technical Report No. 5* (pp. 119–163). University of Hawaii Press.

Clark, H. H., & Gerrig, R. J. (1984). On the pretense theory of irony. *Journal of experimental irony: General*, 113, 121–126.

Colston, H. L. (2000). On necessary conditions for verbal irony comprehension. *Pragmatics and Cognition*, 8, 277–324.

Colston, H. L. (2002). Contrast and assimilation in verbal irony. *Journal of pragmatics*, 34, 111–142.

Colston, H. L. (2017). Irony and sarcasm. In S. Attardo (Ed.), *The Routledge handbook of language and humor* (pp. 234–249). New York: Routledge.

Colston, H. L., & Gibbs, Jr., R. W. (2007). A brief history of irony. In H.L. Colston & R. W. Gibbs, Jr. (Eds.), *Irony in language and thought: A cognitive science reader* (pp. 3–21). Lawrence Earlbaum Associates. Taylor and Francis Group.

Colston, H. L., & O' Brien, J. (2000). Contrast and pragmatics in figurative language: Anything understatement can do, irony can do better. *Journal of Pragmatics* 32, 1557–1593.

Cook, M., & Liddicoat, A. (2002). The development of comprehension in interlanguage pragmatics: The case of request strategies in English. *Australian Review of Applied Linguistics*, 25, 19–39.

Currie, G. (2008). Echo et feintise : quelle est la différence et qui a raison ? *Philosophiques*, 351, 13–23.

Cutler, A. (1974). On saying what you mean without meaning what you say. In M. Galy, R. Fox, & A. Bruck (Eds.), *Papers from the Tenth Regional Meeting, Chicago Linguistic Society* (pp. 117–127). Chicago Linguistic Society.

Cutler, A. (1976). Beyond parsing and lexical look-up: An enriched description of auditory sentence comprehension. In R. Wales & E. Walker (Eds.), *New approaches to language mechanisms: a collection of psycholinguistic studies* (pp. 133–150). North-Holland.

- Davis, W. A. (1998). *Implicature: Intention, Convention, and Principle in the Failure of Gricean Theory*. Cambridge University Press.
- Davis, W. A. (2019). Implicature. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition) URL = <https://plato.stanford.edu/archives/fall2019/entries/implicature> .
- Dews, S., Kaplan, J., & Winner, E. (1995). Why not say it directly? The social functions of irony. *Discourse Processes* 19, 347–367.
- Dews, S., & Winner, E. (1995). Muting the meaning: A social function of irony. *Metaphor and symbolic activity*, 10(1), 3–20.
- Dews, S., Winner, E., Kaplan, J., Rosenblatt, E., Hunt, M., Lim, K., McGovern, A., Qualter, A., & Smarsh, B. (1996). Children's understanding of the meaning and functions of verbal irony. *Child development*, 67(6), 3071–3085.
- D'Imperio, M., Champagne-Lavau, M., Leovenbruck, H., & Fouillaud, L. (2013). Perceptual evaluation of tonal and contextual cues to sarcasm in French. *Phonetics and Phonology in Iberia Conference (PaPI)*. 132–134.
- Dumarsais, C. C. (1730) *Des tropes*. F. Douay-Soublin (Ed.). Flammarion.
- Eggs, E. (2009). Rhétorique et argumentation : de l'ironie. *Argumentation et Analyse du Discours* [Online], 2 URL : <http://aad.revues.org/219>.
- Ervin-Tripp, S., Strage, A., Lampert, M., & Bell, N. (1987). Understanding requests. *Linguistics*, 25(1), 107–144.
- Felix-Brasdefer, J. C. (2008). Pedagogical Intervention and the Development of Pragmatic Competence in Learning Spanish as a Foreign Language. *Issues in Applied Linguistics*, 16(1), 49–84.
- Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage.
- Field, A. P., & Wilcox, R. R. (2017). Robust statistical methods: A primer for clinical psychology and experimental psychopathology researchers. *Behaviour Research and Therapy*, 98, 19–38.
- Field, J. (1998). Skills and strategies: towards a new methodology for listening. *ELT Journal*, 52(2), 110–118.
- Field, J. (2008), *Listening in the Language Classroom*. Cambridge University Press.
- Filik, R., & Moxey, L.M. (2010). The on-line processing of written irony. *Cognition*, 116, 421–436.
- Garcia, P. (2004). Developmental Differences in Speech Act Recognition: A Pragmatic Awareness Study. *Language Awareness*, 13(2), 96–115.

- Gibbs, Jr., R.W. (1979). Contextual effects in understanding indirect requests. *Discourse Processes*, 2, 2–10.
- Gibbs, Jr., R.W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory and cognition*, 8, 449–456.
- Gibbs, Jr., R. W. (1983). Do people always process the literal meanings of indirect requests? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(3), 524–533.
- Gibbs, Jr., R.W. (1986). On the psycholinguistics of sarcasm. *Journal of experimental psychology: General*, 115(1), 3–15.
- Gibbs, Jr., R.W. (1993). Process and products in making sense of tropes. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 252–274). Cambridge University Press.
- Gibbs, Jr., R.W. (1994). Figurative thought and figurative language. In M. A. Gernsbacher (Ed.), *Hand-book of psycholinguistics* (pp. 411–446). Academic.
- Gibbs, Jr., R.W. (2000). Irony in talk among friends. *Metaphor and symbol*, 15(1&2), 5–27.
- Gibbs, Jr., R.W. (2002). A new look at literal meaning in understanding what is said and implicated. *Journal of Pragmatics*, 34, 457–486.
- Gibbs, Jr., R.W., & Colston, H. L. (2012). *Interpreting Figurative Meaning*. Cambridge University Press.
- Gibbs, Jr., R.W., Bryant, Q. A., & Colston, H. L. (2014). Where is the humor in verbal irony? *Humor*, 24(4), 575–595.
- Gibbs, Jr., R.W., & O'Brien, J. E. (1991). Psychological aspects of irony understanding. *Journal of Pragmatics*, 16, 523–530.
- Gibbs, Jr., R.W., O'Brien, J. E., & Doolittle, S. (1995). Inferring meanings that are not intended: Speakers' intentions and irony comprehension. *Discourse Processes*, 20, 187–203.
- Giora, R. (1995). On irony and negation. *Discourse processes*, 19, 239–264.
- Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 7, 183–206.
- Giora, R. (2003). *On our mind: Salience, context and figurative language*. Oxford University Press.
- Giora, R., & Fein, O. (1999). Irony: context and salience. *Metaphor and symbol*, 14, 241–257.
- Giora, R., Fein, O., Laadan, D., Wolfson, J., Zeituny, M., Kidron, R., Kaufman, R., & Shaham, R. (2007). Expecting irony: Context versus salience-based effects. *Metaphor and Symbol*, 22, 119–146.

- Giora, R., Fein, O., & Schwartz, T. (1998). Irony: Graded Salience and Indirect Negation. *Metaphor and Symbol, 13*(2), 83–101.
- Giora, R., Zaidel, E., Soroker, N., Batori, G., & Kasher, A. (2000). Differential Effects of Right- and Left-Hemisphere Damage on Understanding Sarcasm and Metaphor. *Metaphor and Symbol, 15*(1&2), 63–83.
- Glucksberg, S. (1995). Commentary on nonliteral language: Processing and use. *Metaphor and Symbolic Activity, 10*, 1, 47–57.
- Goffman, E. (1955). On face-work: an analysis of ritual elements in social interaction. *Psychiatry: Journal for the Study of Interpersonal Processes, 18*, 213–231.
- Graham, S. (2017). Research into practice: listening strategies in and instructed classroom setting. *Language teaching, 50*(1), 107–119.
- Graham, S., & Macaro, E. (2008). Strategy instruction in listening for lower intermediate learners of French. *Language Learning, 58*(4), 747–783.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41–58). Academic Press.
- Grice, H. P. (1989). *Studies in the way of words*. Harvard University Press.
- Gruba, P. (2005). *Developing media literacy in the L2 classroom*. Macquarie University, National Centre for English Teaching and Research.
- Guidi, A. (2017). Humor Universals. In S. Attardo, (Ed.), *The Routledge Handbook of Language and Humor* (pp. 17–33). Routledge.
- Haiman, J. (1989). Alienation in grammar. *Studies in Language 13*(1), 129–170.
- Haiman, J. (1998). *Talk Is Cheap: Sarcasm, Alienation, and the Evolution of Language*. Oxford University Press.
- Hamamoto, H. (1998). Irony from a cognitive perspective. In R. Carston & S. Uchida (Eds.) *Relevance theory. Applications and implications* (pp. 257–270). Amsterdam: John Benjamins.
- Hamblin, J., & Gibbs, R. (2003). Processing the meaning of what speakers say and implicate. *Discourse Processes, 35*, 59–80.
- Haugh, M. (2017). Teasing. In S. Attardo (Ed.), *The Routledge handbook of language and humor* (pp. 204–218). Routledge.
- Haverkate, H. (1990). A speech act analysis of irony. *Journal of Pragmatics, Vol. 14*(1), 77–109.
- Holtgraves, T. (1999). Comprehending indirect replies: When and how are their conveyed meanings activated? *Journal of Memory and Language, 38*, 519–540.

- Horn, L. R. (1984). Towards a New Taxonomy for Pragmatic Inference: Q-based and R-based Implicature. In D. Schiffrin (Ed.), *Georgetown University Round Table on Languages and Linguistics* (pp. 11–42). Georgetown University Press.
- Horn, L. R. (2004). Implicature. In L. R. Horn & G. Ward (Eds.), *The Handbook of Pragmatics* (pp. 2–28). Blackwell.
- In'nami, Y., & Koizumi, R. (2022). The relationship between L2 listening and metacognitive awareness across listening tests and learner samples. *International Journal of Listening*, 36(2), 100–117.
- Ishara, N., & Cohen, A. D. (2010). *Teaching and learning pragmatics: Where language and culture meet*. Longman Applied Linguistics. Pearson Education Limited.
- Ivanko, S. L., & Pexman, P. M. (2003). Context incongruity and irony processing. *Discourse processes*, 35(3), 241–279.
- Jasny, B. R., Chin, G., Chong, L., & Vignieri, S. (2011). Data replication & reproducibility. Again, and again, and again .... Introduction. *Science*, 334(6060), 1225.
- Jobert, M., & Sorlin, S. (2018). Introduction. The intricacies of irony and banter. In M. Jobert and S. Sorlin (Eds.), *The Pragmatics of Irony and Banter* (pp. 3–21). John Benjamins Publishing Company.
- Jorgensen, J. (1996). The functions of sarcastic irony in speech. *Journal of Pragmatics* 26, 613–634.
- Jorgensen, J., Miller, G. A., & Sperber, D. (1984). Test of the mention theory of irony. *Journal of experimental psychology*, 113(1), 112–120.
- Kapogianni, E. (2014). Differences in use and function of verbal irony. *Humor: International Journal of Humor Research*, 27, 597–618.
- Kasper, G. (1984). Pragmatic comprehension in learner-native speaker discourse. *Language Learning*, 34, 1–20.
- Katz, A. N. (2000). Introduction to the special issue: the uses and processing of irony and sarcasm. *Metaphor and symbol*, 15(1&2), 1–3.
- Kaufert, D. S., 1981. Understanding ironic communication. *Journal of Pragmatics*, 5, 495–510.
- Keenan, E. O. (1976). The universality of conversational postulates. *Language in Society*, 5, 1, 67–80.
- Kerbrat-Orecchioni, C. (1976). Problèmes de l'ironie. *Linguistique et sémiologie*, 2, 10–47.
- Kerbrat-Orecchioni, C. (1980). L'ironie comme trope. *Poétique*, 41, 108–127.
- Kerbrat-Orecchioni, C. (1992). *Les interactions verbales*. Armand Colin.

- Kerbrat-Orecchioni, C. (2008). *Les Actes de Langage Dans le Discours*. Armand Colin.
- Kida, I. (2011). Language distance across cultures as a way of expressing politeness and not only. In J. Arabski & A. Wojtaszek (Eds.) *Aspects of Culture in Second Language Acquisition and Foreign Language Learning* (pp. 183–191). Berlin Heidelberg.
- Kim, J. (2013). *Developing conceptual understanding of sarcasm in a second language through concept-based instruction*. Unpublished dissertation in Applied Linguistics. Pennsylvania State University.
- Kim, J. (2014). How Korean EFL learners understand sarcasm in L2 English. *Journal of Pragmatics*, 60, 193–206.
- Kim, J. (2016). Teaching language learners how to understand sarcasm in L2 English. In N. Bell (Ed.), *Multiple perspectives on language play* (pp.317–346). De Gruyter Mouton.
- Kim, J., & Lantolf, J. P. (2016). Developing conceptual understanding of sarcasm in L2 English through explicit instruction. *Language Teaching Research*, 1(22), 1–22.
- Koike, D. A. (1996). Transfer of pragmatic competence and suggestions in Spanish foreign language learning. In S. M. Gass & J. Neu (Eds.), *Speech Acts Across Cultures: Challenges to Communication in a Second Language* (pp. 257–281). Mouton de Gruyter.
- Kondo, S. (2010). Apologies: Raising learners' cross-cultural awareness. In A. Martinez-Flor & E. Uso-Juan (Eds.), *Speech act performance: theoretical, empirical and methodological issues* (Language Learning & Language Teaching) (pp. 145–162). John Benjamins Publishing Company.
- Kreuz, R. J. (1996). The use of verbal irony: Cues and constraints. In J. Mio & A. Katz (Eds.), *Metaphor: Pragmatics and applications* (pp. 23–38). Lawrence Erlbaum Associates.
- Kreuz, R. J. (2000). The production and processing of verbal irony. *Metaphor and symbol*, 15(1&2), 99–107.
- Kreuz, R. J., & Glucksberg, S. (1989). How to be sarcastic: The echoic reminder theory of verbal irony. *Journal of Experimental Psychology*, 118(4), 374–386.
- Kreuz, R. J., Long, D. L., & Church, M. B. (1991). On being ironic: Pragmatic and mnemonic implications. *Metaphor and symbolic activity*, 6(3), 149–162.
- Kreuz, R. J., & Roberts, R. M. (1995). Two cues for verbal irony: Hyperbole and the ironic tone of voice. *Metaphor and Symbolic Activity*, 10(1), 21–31.
- Krulatz, A., & Dixon, T. (2020). The use of refusal strategies in interlanguage speech act performance of Korean and Norwegian users of English. *Studies in Second Language Learning and Teaching*, 10(4), 751–777.
- Kubota, M. (1995). Teachability of Conversational Implicature to Japanese EFL Learners. *IRLT (Institute for Research in Language Teaching) Bulletin*, 9, 35–67.



- Kumon-Nakamura, S., Glucksberg, S., & Brown, M. (1995). How about another piece of pie: The allusional pretense theory of discourse irony. *Journal of experimental psychology*, 124(1), 3–21
- Lachman, R., Lachman, J. L., & Butterfield, E. C. (1979). *Cognitive psychology and information processing: An introduction*. Lawrence Erlbaum Associates.
- Lakoff, R. T. (1973). The logic of politeness; or, minding your P's and Q's. In C. Corum, T. Cedric Smith-Stark, & A. Weiser (Eds.), *Papers from the Ninth Regional Meeting of the Chicago Linguistics Society* (pp. 292–305). University of Chicago.
- Lakoff, R. T. (1990). *Talking power. The politics of language*. Basic Books/Harper Collins.
- Larking-Galiñanes, C. (2017). An overview of humor theory. In S. Attardo (Ed.), *The Routledge handbook of language and humor* (pp.4–16). Routledge
- Leech, G. (1983). *Principles of Pragmatics*. Longmans.
- Leech, G. (2014). *The Pragmatics of Politeness*. Oxford University Press.
- Leech, G. (2005). Politeness: is there an East-West divide. *Journal of Foreign Languages*, 6(3), 3–31.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.
- Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge MIT Press.
- Levorato, M. C. (1993). The acquisition of idioms and the development of figurative competence. In C. Cacciari & P. Tabossi (Eds.), *Idioms: Processing, structure, and interpretation* (pp. 101–128). Lawrence Erlbaum Associates.
- Maeshiba, N., Yoshinaga, N., Kasper, G., & Ross, S. (2006). Transfer and proficiency in interlanguage apologizing. In S. M. Gass & J. Neu (Eds.), *Speech Acts Across Cultures: Challenges to Communication in a Second Language* (pp. 155–190). Mouton de Gruyter.
- Mair, P., & Wilcox, R. R. (2020). Robust Statistical Methods in R Using the WRS2 Package. *Behavior Research Methods*, 52, 464–488.
- Manno, G. (1999). Savoir refuser à l'écrit : analyse d'un enchaînement non préféré de macro-actes de discours. *Journal of French Language Studies*, 9(1), 39–68.
- Manno, G. (2002). La politesse et l'indirection : un essai de synthèse. *Langage et société*, 100, 5–47.
- Markee, N. (2017). Are Replication Studies Possible in Qualitative Second/Foreign Language Classroom Research? A Call for Comparative Re-Production Research. *Language Teaching*, 50(3), 367–383.

- Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018a). Replication in Second Language Research: Narrative and Systematic Reviews and Recommendations for the Field. *Language Learning*, 68(2), 321–391.
- Marsden, E., Morgan-Short, K., Trofimovich, P., & Ellis, N. C. (2018b). Introducing Registered Reports at Language Learning: Promoting Transparency, Replication, and a Synthetic Ethic in the Language Sciences. *Language Learning*, 68(2), 309–320.
- Martinez-Flor, A., & Uso-Juan, E. (2010). Pragmatics and speech act performance. In A. Martinez-Flor & E. Uso-Juan (Eds.), *Speech act performance: theoretical, empirical and methodological issues* (Language Learning & Language Teaching) (pp. 3–20). John Benjamins Publishing Company.
- Mbow, F. (2011). L'acte de refus et le fonctionnement de la politesse dans les rencontres commerciales. *LIENS*, 14, 181–206.
- McManus, K. (2021). Replication and Open Science in Applied Linguistics Research. In L. Plonsky (Ed.), *Open Science in Applied Linguistics*. John Benjamins. (OSF Preprint) <https://doi.org/10.31219/osf.io/bqr9w>.
- McManus, K., & Marsden, E. (2018). Online and offline effect of L1 practice in L2 grammar learning: a partial replication. *Studies in Second Language Acquisition*, 40(2), 459–475.
- Meisner, C., Robert-Tissot, A., & Stark, E., (2015). L'absence et la présence du NE de négation, in *Encyclopédie Grammaticale du Français*: <http://encyclogram.fr>.
- Mendelsohn, D. (1995). Applying Learning Strategies in the Second/Foreign Language Listening Comprehension Lesson. In D. Mendelsohn & J. Rubin (Eds.) *A Guide for the Teaching of Second Language Listening* (pp. 132–150). Dominic Press, Inc.
- Milanowicz, A. (2013). Irony as a Means of Perception Through Communication Channels. Emotions, Attitude and IQ Related to Irony Across Gender. *Psychology of Language and Communication*, 17(2), 115–132.
- Muecke, D. C. (1978). Irony markers. *Poetics*, 7(4), 363–375.
- Mulo Farenkia, B. (2015). Invitation refusals in Cameroon French and Hexagonal French. *Multilingua - Journal of Cross-Cultural and Interlanguage Communication*, 34(4), 577–603.
- Mulo Farenkia, B. (2019). Offer Refusals in L2 French. In C. Savvidou (Ed.), *Second Language Acquisition - Pedagogies, Practices and Perspectives*. IntchOpen <https://doi.org/10.5772/intechopen.87189>.
- Mulo Farenkia, B. (2020). Refuser une demande de service en français L1 et L2. *European Journal of Applied Linguistics Studies*, 2(2), 30–50.
- Muyskens, J., Harlow, L., Vialet, M., & Brière, J-F. (2015). *Bravo!* 8th ed. Heinle & Heinle.
- Myers Roy, A. (1981). The function of irony in discourse. *Text* 1(4), 407–423.

- Oomen, C. C. E., & Postma, A. (2001). Effects of time pressure on mechanisms of speech production and self-monitoring. *Journal of Psycholinguistic Research*, 30, 163–184.
- Oxford, R. (1990). *Language Learning Strategies*. Heinle and Heinle.
- Paesani, K., Allen, H. W., & Dupuy, B. (2016). *A multiliteracies framework for collegiate foreign language teaching*. Pearson
- Paltridge, B. (2006). *Discourse Analysis*. Continuum.
- Paris, J. (2015). *Lumière sur le développement de la production de langage non-littéral en L2. Pour une comparaison avec l'acquisition des langues maternelles*. Doctoral Dissertation : English Linguistics. Jury: Morgenstern, A., Hilton, H., Tardieu-Garnier, C., Jamet, D., Littlemore, J., Parris, C. Sorbonne Paris Cité.
- Partridge, C., Moberg, M. (Eds.) (2017). *The Bloomsbury Handbook of Religion and Popular Music*. Bloomsbury Publishing.
- Plonsky, L. (2015). *Advancing quantitative methods in second language research*. Routledge.
- Plonsky, L., & Oswald, F. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912.
- Porte, G. K. (2012). *Replication research in applied linguistics*. Cambridge University Press.
- Porte, G. K., & McManus, K. (2019). *Doing replication research in applied linguistics*. Routledge.
- Porte, G. K., & Richards, K. (2012). Focus article: Replication in second language writing research. *Journal of Second Language Writing*, 21(3), 284–293.
- Regel, S., Gunter, T.C., & Friederici, A. D. (2011). Isn't it ironic? An electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience*, 23, 77–293.
- Ren, W. (2015). Sociopragmatic Variation in Mainland and Taiwan Chinese Refusals. In K. Beeching & H. Woodfield (Eds.), *Researching Sociopragmatic Variability* (pp. 72–93). Palgrave Macmillan.
- Roberts, R. M., & Kreuz, R. J. (1994). Why do people use figurative language? *Psychological Science*, 5(3), 159–163.
- Roever, C. (2005). *Testing EFL pragmatics*. Gunter Narr.
- Rose, K. R. (2005). On the effects of instruction in second language pragmatics. *System*, 33, 385–399.
- Rose, K. R., & Kasper, G. (Eds.). (2001). *Pragmatics in language teaching*. Cambridge University Press.

- Ross, S. (1997). An introspective analysis of listener inferencing on a second language listening test. In G. Kasper & E. Kellerman (Eds.), *Communication strategies: Psycholinguistic and sociolinguistic perspectives* (pp. 216–237). Longman.
- Rost, M., & Ross, S. (1991). Learner use of strategies in interaction: Typology and teachability. *Language Learning*, 41(2), 235–273.
- Roussel, S., & Tricot, A. (2012). A tentative approach to analysing listening strategies in CALL. *Procedia - Social and Behavioral Sciences*, 34, 193–197.
- Rubin, J. (1995). An overview to A guide for the teaching of second language listening. In D. J. Mendelsohn & J. Rubin (Eds.), *A guide for the teaching of second language listening* (pp. 7–11). Dominie Press.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50(4), 696–735.
- Schaffer, Rachel R., 1982. *Vocal clues for irony in English*. Unpublished Ph.D. dissertation, Ohio State University.
- Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge University Press.
- Schegloff, E. A., & Sacks, H. (1974) Opening up Closings. *Semiotica*, 8(4), 289–327.
- Schmidt, R. (1998). Consciousness, learning, and interlanguage pragmatics. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage Pragmatics* (pp. 21–42). Oxford University Press.
- Schwoebel, J., Dews, S., Winner, E., & Srinivas, K. (2000). Obligatory processing of the literal meaning of ironic utterances: Further evidence. *Metaphor and symbol*, 15(1&2), 47–61.
- Searle, J. R. (1975). Indirect speech acts. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (pp. 59–82). Academic Press.
- Searle, J. R., & Vanderveken, D. (2009). *Foundations of illocutionary logic*. Cambridge University Press.
- Seto, K.-I. (1998). On non-echoic irony. In R. Carston & S. Uchida (Eds.), *Relevance theory. Applications and implications* (pp. 239–256). John Benjamins.
- Shapely, M. (1987). Prosodic variation and audience response. *IprA: Papers in Pragmatics* 1(2), 66–79
- Shively, R. L. (2011). L2 Pragmatic development in study abroad: A longitudinal study of Spanish service encounters. *Journal of Pragmatics*, 43, 1818–1835.
- Shively, R. L. (2013). Learning to be funny in Spanish during study abroad: L2 humor development. *The modern language journal*, 97(4), 930–946.
- Shively, R. L., Menke, M. R., & Manzon-Omundson, S. M. (2008). Perception of irony by L2 learners of Spanish. *Issues in applied linguistics*, 16(12), 101–132.

- Sperber, D., & Wilson, D. (1981). Irony and use-mention distinction. In P. Cole (Ed.), *Radical Pragmatics*, (pp. 295–318). Academic Press.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Basil Blackwell.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2<sup>nd</sup> edition). Blackwell.
- Sperber, D., & Wilson, D. (1998). Irony and relevance: A reply to Drs. Seto, Hamamoto and Yamanashi. In R. Carston & S. Uchida (Eds.), *Relevance theory: Applications and implications* (pp. 283–293). John Benjamins Publishing Company.
- Sperber, D., & Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind and Language*, 17, 3–23.
- Stavans, A., & Shafran, R. W. (2018). The pragmatics of requests and refusals in multilingual settings. *International journal of multilingualism*, 15(2), 149–168.
- Taguchi, N. (2002). An application of relevance theory to the analysis of L2 interpretation processes: The comprehension of indirect replies. *International Review of Applied Linguistics*, 40, 151–176.
- Taguchi, N. (2005). Comprehension of implied meaning in English as a second language. *Modern Language Journal*, 89, 543–562.
- Taguchi, N. (2007). Development of Speed and Accuracy in Pragmatic Comprehension in English as a Foreign Language. *Tesol Quarterly: a Journal for Teachers of English to Speakers of Other Languages and of Standard English As a Second Dialect*, 41(2), 313–338.
- Taguchi, N. (2008a). Pragmatic comprehension in Japanese as a foreign language. *Modern Language Journal*, 92, 558–576.
- Taguchi, N. (2008b). Cognition, language contact, and development of pragmatic comprehension in a study-abroad context. *Language Learning*, 58, 33–71.
- Taguchi, N. (2008c). The role of learning environment in the development of pragmatic comprehension. A comparison of Gains Between EFL and ESL Learners. *Studies in Second Language Acquisition*, 30(4), 423–452.
- Taguchi, N. (2009a). Comprehension of indirect opinions and refusals in L2 Japanese. In N. Taguchi (Ed.), *Pragmatic Competence* (pp. 249–247). Mouton de Gruyter.
- Taguchi, N. (2009b). Corpus-informed assessment of comprehension of conversational implicatures in L2 English. *TESOL Quarterly*, 43(4), 738–749.
- Taguchi, N. (2011). The effect of L2 proficiency and study-abroad experience in pragmatic comprehension. *Language Learning*, 61, 904–939.

- Taguchi, N. (2012). *Context, individual differences, and pragmatic competence*. Multilingual Matters.
- Taguchi, N. (2019). *The Routledge handbook of second language acquisition and pragmatics*. Routledge.
- Taguchi, N., Gomez-Laich, M. P., & Arrufat-Marques, M.-J. (2016). Comprehension of Indirect Meaning in Spanish as a Foreign Language. *Foreign Language Annals*, 49(4), 677–698.
- Taguchi, N., Li, S., & Liu, Y. (2013). Comprehension of conversational implicature in L2 Chinese. *Pragmatics & Cognition*, 21(1), 139–157.
- Taguchi, N., & Yamaguchi, S. (2019). Implicature Comprehension in L2 Pragmatics Research. In N. Taguchi (Ed.), *The Routledge handbook of second language acquisition and pragmatics* (pp. 31–46). Routledge.
- Takahashi, S. (1990). Exploring the Comprehension Process of Nonliteral Utterances and Some Implications for Automaticity. *University of Hawai'i Working Papers in ESL*, 9(2), 67–97.
- Takahashi, S. (2010). The effect of pragmatic instruction on speech act performance. In A. Martinez-Flor & E. Uso-Juan (Eds.), *Speech act performance: theoretical, empirical and methodological issues* (Language Learning & Language Teaching) (pp. 127–142). John Benjamins Publishing Company.
- Takahashi, S. & Beebe, L. (1987). The development of pragmatic competence by Japanese learners of English. *Japan Association of Language Teachers (JALT) Journal*, 8, 131–155.
- Takahashi, S., & Roitblat, H. L. (1994). Comprehension Process of Second Language Indirect Requests. *Applied Psycholinguistics*, 15(4), 475–506.
- Tanck, S. (2002). Speech act sets of refusal and complaint: A comparison of native and non-native English speakers' production. *TESOL Working Papers*.
- Tannen, D. (1987). Conversational style. In H. W. Dechert & M. Raupach (Eds.), *Psycholinguistic Models of Production* (pp. 251–267). Ablex.
- Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics*, 4, 91–112.
- Thomas, J. (1995). *Meaning in interaction: An introduction to pragmatics*. Longman.
- Vandergrift, L. (2003a). Orchestrating strategy use: Toward a model of the skilled second language/Listener. *Language Learning*, 53(3), 463–496.
- Vandergrift, L. (2003b). From prediction through reflection: Guiding students through the process of L2 listening. *Canadian Modern Language Review*, 59, 425–440.
- Vandergrift, L. (2006). Second Language Listening: Listening Ability or Language Proficiency? *The Modern Language Journal*, 90(1), 6–18.

- Vandergrift, L., & Goh, C. (2012). *Teaching and learning second language listening - Metacognition in action*. Taylor & Francis.
- VanPatten, B. (Ed.) (2004). *Processing instruction: theory, research, and commentary*. Lawrence Erlbaum Associates.
- Veenman, M. V. J. (2011). Alternative assessment of strategy use with self-report instruments: a discussion. *Metacognition Learning* 6, 205–211.
- Verschueren, J. (1999). *Understanding pragmatics*. Oxford University Press.
- Wallmark, J. M. (1971). *Sarcasm: An analysis of and suggestions for teaching the recognition of sarcasm to non-native speakers*. M.A. thesis. The Library of the University of California, Los Angeles.
- Weizman, E. (1985). Towards an analysis of opaque utterances: hints as a request strategy. *Theoretical Linguistics* 12(2&3), 153–63.
- Wilcox, R. (2012). *Introduction to Robust Estimation and Hypothesis Testing* (3rd ed.). Elsevier.
- Wilson, D. (2006). The pragmatics of verbal irony: Echo or pretense? *Lingua*, 116(10), 1722–1743.
- Wilson, D. (2013). Irony comprehension: A developmental perspective. *Journal of Pragmatics*, 59, 40–56.
- Wilson, D., & Sperber, D. (1992). On verbal irony. *Lingua*, 87, 53–76.
- Wilson, D., & Sperber, D. (1998). Pragmatics and time. In R. Carston & S. Uchida (Eds.), *Relevance Theory*, [Pragmatics and beyond new series, 37] (pp. 1–22). John Benjamins Publishing Company.
- Wilson, D., & Sperber, D. (2002). Relevance theory. *UCL Working Papers in Linguistics*, 14, 249-287.
- Wilson, D., & Sperber, D. (2007). On verbal irony. In H. L. Colston & R. W. Gibbs, Jr. (Eds.), *A brief history of irony. In Irony in language and thought: A cognitive science reader* (pp. 35–55). Lawrence Earlbaum Associates.
- Wilson, D., & Sperber, D. (2012). *Meaning and Relevance*. Cambridge University Press.
- Yamanaka, J. E. (2003). Effects of proficiency and length of residence on the pragmatic comprehension of Japanese ESL learners. *Second Language Studies*, 22(1), 107–175.
- Yamanashi, M. (1998). Some issues in the treatment of irony and related tropes. In R. Carston & S. Uchida (Eds.), *Relevance Theory: Applications and Implications* (pp. 271–280). John Benjamins.
- Yus, F. (1998). Irony: context accessibility and processing effort. *Pragmalingüística* 5(6), 391–411.

Yus, F. (2000). On reaching the intended ironic interpretation. *International Journal of Communication* 10(1&2), 27–78.

Yus, F. (2017). Relevance-theoretic treatments of humor. In S. Attardo (Ed.), *The Routledge handbook of language and humor* (pp. 189–203). Routledge.

Zhang, Y. (2012). The Impact of Listening Strategy on Listening Comprehension. *Theory and Practice in Language Studies*, 2(3), 625–629.

Zoghalmi, N. (2016). La compréhension de l'anglais oral (l2) : processus cognitifs et comportements stratégiques. *Recherche et pratiques pédagogiques en langues de spécialité, [En ligne]*, 35(1).

Zoghalmi, N., & Hilton, H. E. (2021). La compréhension de l'oral en langue étrangère. In P. Leclercq, A. Edmonds, & E. Sneed German (Eds.), *Introduction à l'acquisition des langues étrangères* (pp. 145–161). DeBoeck Supérieur.



# APPENDICES

## Appendix A. Summary of Taguchi's studies on L2 implicature comprehension

**Table 25**

*Summary of Taguchi's studies on L2 implicature comprehension*

Year	Major findings
2002	<ul style="list-style-type: none"> <li>- Significant difference between lower proficiency and higher proficiency learners in comprehension of indirect meaning (indirect opinions, indirect refusals, disclosures).</li> <li>- Verbal reports showed that lower proficiency learners could reach comprehension but had lower confidence level</li> <li>- Disclosures were harder than indirect opinions.</li> <li>- Indirect refusals were easier due to conventional pattern</li> <li>- Use of multiple inferential strategies when implicature is more opaque (less conventional).</li> <li>- Higher proficiency learners used more paralinguistic cues (to understand attitudinal and emotional messages) and adjacency pair rule (for indirect opinions).</li> <li>- Lower proficiency learners used more background knowledge/experience and keyword inferencing.</li> </ul>
2005	<ul style="list-style-type: none"> <li>- No difference in native speakers' comprehension between More Conventional Implicatures (MCI) and Less Conventional Implicatures (LCI).</li> <li>- LCIs (indirect opinions) were harder than MCIs (indirect requests and refusals)</li> <li>- Proficiency impacted accuracy for both MCIs and LCIs and speed for MCI (to a smaller extent).</li> <li>- No speed-accuracy trade-off: response time was higher on incorrect items.</li> <li>- Analysis of wrong answers/distractors: learners were most attracted to the third option based on "words taken from the last part of the dialogue" (p. 550), especially for the LCIs. → recency effect</li> <li>- No significant difference between L2 accuracy and speed.</li> <li>- No significant relationship between L1 comprehension speed and L2 comprehension speed.</li> </ul>
2007	<ul style="list-style-type: none"> <li>- NS were more accurate and faster than L2 learners.</li> <li>- L2 learners' accuracy and speed for indirect refusals were higher than for indirect opinions.</li> <li>- Significant overall gains over 7-week period (more for accuracy than for speed)</li> <li>- Lexical access speed correlates with response time</li> <li>- Proficiency correlates with accuracy</li> </ul>
2008a	<ul style="list-style-type: none"> <li>- Overall indirect refusals were easier than indirect opinions (L2 Japanese)</li> <li>- Significant difference between elementary and intermediate proficiency levels</li> <li>- Conventional indirect opinions were the most difficult for intermediate proficiency group.</li> <li>- Both types of indirect opinions were hard for the elementary proficiency group.</li> <li>- Comprehension was faster for indirect refusals and slower for nonconventional indirect opinions for both groups.</li> <li>- Significant difference between pragmatic processing and basic processing speed was found</li> </ul>

	<ul style="list-style-type: none"> <li>- Analysis of wrong answers/distractors: intermediate learners were most attracted to the third option based on “words taken from the last heard utterance of the dialogue or their synonyms” (p. 566) → recency effect. elementary learners were attracted to both the first (“meaning opposite to the implied meaning” p. 566) and second option → both recency effect and poor listening proficiency</li> <li>- Verbal reports showed listening comprehension difficulties for elementary level</li> <li>- Impact of knowledge pragmalinguistic devices (more proficiency learners were better)</li> </ul>
2008b	<ul style="list-style-type: none"> <li>- ESL learners consistently comprehended conventional indirect refusals more accurately and faster than nonconventional indirect opinions (T1, T2, T3)</li> <li>- Significant gains in accuracy for indirect refusals between T1 and T3.</li> <li>- No gain for indirect opinions</li> <li>- Significant gain in speed for indirect refusals (all contrasts for T1, T2, T3) and for indirect opinion between (except for T2 and T3)</li> <li>- Lexical judgment improved over time and correlated with gains in speed</li> <li>- Lexical judgment gain did not correlate with accuracy gains</li> <li>- Amount of reported L2 language contact did not correlate with gains in accuracy</li> <li>- Amount of reported L2 speaking and reading correlated with gains in speed.</li> <li>- Amount of reported L2 listening and reading did not correlate with gains in speed.</li> </ul>
2008c	<ul style="list-style-type: none"> <li>- Significant gains for both ESL and EFL groups in accuracy and speed of comprehension of indirect refusals and indirect opinions.</li> <li>- Gains were more significant in speed for ESL and in accuracy for EFL</li> <li>- EFL gains were more significant for indirect refusals</li> <li>- Overall comprehension was easier for indirect refusals due to conventionality</li> <li>- Gains did not correlate with the reported amount of speaking</li> </ul>
2009a	<ul style="list-style-type: none"> <li>- Overall indirect refusals were easier than indirect opinions (L2 Japanese)</li> <li>- Conventional indirect opinions were the most difficult.</li> <li>- Significant difference between elementary and intermediate+advanced proficiency levels</li> <li>- Verbal reports showed listening comprehension difficulties for elementary level</li> <li>- Impact of knowledge of pragmalinguistic devices for more advanced learners but potential for negative L1 transfer</li> </ul>
2009b	<ul style="list-style-type: none"> <li>- NS were more accurate and faster than L2 learners. NS speed was uniform across three types of indirect meaning.</li> <li>- Indirect refusals were easier to understand due to conventionality</li> <li>- Significant difference in accuracy between indirect refusals and less conventional implicatures (indirect opinion or comments) and between indirect refusals and routines.</li> <li>- Refusals took less time and less conventional implicatures more time to comprehend</li> <li>- Gains in accuracy for refusals only (no significant gains in speed)</li> </ul>
2011	<ul style="list-style-type: none"> <li>- NS were more accurate and faster than L2 learners. NS speed was uniform across three types of indirect meaning.</li> <li>- Conventional implicatures (indirect refusals and routines) were easier and faster to comprehend than nonconventional implicatures (indirect comments and opinions)</li> <li>- Indirect refusals were easier/faster overall (due to conventionality)</li> <li>- Proficiency and study abroad experience impacted comprehension accuracy of nonconventional implicature</li> <li>- Proficiency (but not study abroad) impacted comprehension accuracy of conventional implicature</li> <li>- Within conventional implicatures: proficiency and study abroad impacted accuracy for routines but only proficiency impacted indirect refusal accuracy.</li> <li>- Impact of proficiency (but not study abroad) for comprehension speed (except for routines that did not have any difference).</li> </ul>

## Appendix B. Sample Comparison of Original and Modified Text in Multimedia Listening Test for Indirect Refusal, from Taguchi et al. (2013) – Appendix – p. 698

### Original Text

*Antonio: Hola, Margarita. Recién te pedí una chocolatada caliente. Hace un frío de morirse.*

*Margarita: Sos, eh. No era necesario.*

*Antonio: ¿Quieres algo para picar? No pedí nada porque tu gusto es un tanto complicado.*

*Margarita: Piqué algo en casa.*

[Antonio: Hi, Margarita! I ordered something for you to drink. I ordered you a hot chocolate. We are gonna perish from this cold weather.

Margarita: Dude, that wasn't necessary.

Antonio: Would you like something to nibble on? I didn't order anything because you are weird when it comes to food.

Margarita: I nibbled on something at home.]

### Modified Text

*Antonio: Hola, Margarita. Te acabo de pedir algo para tomar. Te pedí un chocolate caliente. Hace mucho frío hoy.*

*Margarita: Ay, Antonio, ¡muchas gracias! No era necesario.*

*Antonio: ¿Quieres algo de comer? No pedí nada porque no sabía qué te apetecería.*

*Margarita: Almorcé en casa.*

[Antonio: Hi, Margarita! I ordered something for you to drink. I ordered you a hot chocolate. It's so cold today.

Margarita: Thank you, Antonio!

Antonio: Would you like something to eat? I didn't order anything because I didn't know what you'd like.

Margarita: I had lunch at home.]

## Appendix C. Test items (French and English version)

*[item short descriptions that have been crossed out indicate that the item was not retained for the final L2 learner test]*

### 1. Conventional Implicatures: Indirect Refusals and Irony

#### Indirect Refusals (IRF)

##### IRF1repas – Q1

Voici une conversation entre François (homme) et Cécile (femme). Ils font des projets pour ce soir.

François: Salut Cécile! Qu'est-ce que tu vas faire ce soir? Tu veux aller au cinéma?

Cécile: Non, je peux pas. J'ai beaucoup de choses à faire.

François: Tu veux dîner avec moi? Ça ne prendra pas trop de temps.

Cécile: Tu manges pas assez vite.

[41]

1. Cécile ne veut pas dîner avec François ce soir. (CORRECT)
  2. Cécile va dîner avec François ce soir. (Principle 1)
  3. Cécile pense que les serveurs au restaurant vont vite. (Principle 2)
  4. Cécile pense qu'elle a le temps de dîner avec François ce soir. (Principle 3)
- (37)

##### IRF1repas

This is a conversation between François (male speaker) and Cécile (female speaker). They are making plans for tonight.

François: Hi, Cécile! What are you doing tonight? Wanna go to the movies?

Cécile: I can't. I have lots of things to do.

François: Do you want to have dinner with me? It won't take us that long.

Cécile: You don't eat fast enough.

(39)

1. Cécile doesn't want to have dinner with François tonight. (CORRECT)
  2. Cécile is going to have dinner with François tonight. (Principle 1)
  3. Cécile thinks waiters at the restaurant are fast. (Principle 2)
  4. Cécile thinks she has time to have dinner with François tonight. (Principle 3)
- (37)

##### IRF2cafétéria – Q2

Interaction entre deux amis, Fabrice (homme) et Marie (femme). Ils sont dans une cafétéria et discutent de ce qu'ils pourraient prendre.

Fabrice: Salut Marie! Je t'ai commandé quelque chose à boire. J'ai pris un chocolat chaud; il fait si froid aujourd'hui.

Marie: Merci Fabrice!

Fabrice: Est-ce que tu voudrais quelque chose à manger? Je n'ai rien commandé car je ne savais pas ce que tu préfères.

Marie: C'est bon. J'ai déjà déjeuné à la maison.

(50)

1. Marie ne veut rien à manger. (CORRECT)
2. Marie veut manger quelque chose avec son chocolat chaud. (Principe 1)
3. Marie va prendre son chocolat chaud à la maison. (Principe 2)
4. Marie voudrait bien prendre une part de tarte aux pommes avec son chocolat chaud. (Principe 3)

(38)

### **IRF2cafétéria**

This is an interaction between two friends, Fabrice (male speaker) and Marie (female speaker). They are in a cafeteria and they are talking about having something to eat.

Fabrice: Hi, Marie! I ordered something for you to drink. I ordered you a hot chocolate. It's so cold today.

Marie: Thank you, Fabrice!

Fabrice: Would you like something to eat? I didn't order anything because I didn't know what you'd like.

Marie: I'm good. I already had lunch at home.

(47)

5. Marie doesn't want anything to eat. (CORRECT)
6. Marie wants something to eat together with the hot chocolate. (Principe 1)
7. Marie will have her hot chocolate at home. (Principe 2)
8. Marie wants to eat a piece of apple pie together with the hot chocolate. (Principe 3)

(38)

### **IRF3voyage - Q3**

Conversation entre deux amis, Marc (homme) and Natalie (femme). Ils discutent de la possibilité de faire un voyage ensemble.

Marc: Je viens de voir que les billets d'avion pour Londres ne sont pas chers du tout!

Marie: Ah bon? Normalement ils sont très chers. Combien coûte le billet aller-retour?

Marc: Environ 250\$! Est-ce que tu voudrais y aller avec moi?

Marie: Je n'ai même pas l'argent pour m'acheter une veste.

(47)

1. Marie n'a pas assez d'argent pour aller en visite à Londres. (CORRECT)
2. Marie pourra se rendre à Londres avec Marc. (Principe 1)
3. Marie veut acheter une nouvelle veste qui coûte environ 250\$. (Principe 2)
4. Marc et Marie vont déménager à Londres le mois prochain. (Principe 3)

(39)

### **IRF3voyage**

This is a conversation between two friends, Marc (male speaker) and Natalie (female speaker). They are talking about the possibility of going on a trip together.

Marc: Just saw the plane tickets for London are quite inexpensive.

Marie: Really? They are normally expensive. How much is a round trip ticket?

Marc: It's \$250 per person! Would you like to come with me?

Marie: I don't even have money for a jacket.

(41)

1. Marie does not have enough money to go and visit London. (CORRECT)
2. Marie will be able to visit London with Marc. (Principle 1)
3. Marie wants to buy a new jacket which costs \$250. (Principle 2)
4. Marie and Marc are moving to London next month. (Principle 3)

(39)

### **IRF4révisions - Q4**

Conversation entre deux étudiants, Franck (homme) et Sophie (femme). Ils discutent de la possibilité d'étudier ensemble après le cours pour se préparer à leur examen de biologie.

Franck: L'examen de biologie est la semaine prochaine. Il y a des modules que je n'ai pas trop compris.

Sophie: Oui ; il y a des modules très difficiles ce semestre.

Franck: Est-ce que tu voudrais réviser avec moi?

Sophie: Je préférerais étudier seule pour les examens.

(43)

1. Sophie ne va pas étudier avec Franck pour l'examen de biologie. (CORRECT)
2. Sophie et Franck étudieront ensemble pour l'examen de biologie. (Principle 1)
3. Sophie n'aime pas étudier seule avant les examens. (Principle 2)
4. Sophie et Franck vont prendre le même cours de biologie le semestre prochain. (Principle 3)

(41)

### **IRF4révisions**

This is a conversation between two classmates, Franck (male speaker) and Sophie (female speaker). They are talking about studying together for their biology exam after their class.

Franck: Our biology exam is coming soon. There are a couple of modules I don't quite understand.

Sophie: Yes. There are some very difficult modules this semester.

Franck: Would you like to study with me for the exam?

Sophie: I prefer to study alone for my exams.

(43)

1. Sophie is not studying with Franck for their Biology exam. (CORRECT)
2. Sophie and Franck will study together for their Biology exam. (Principle 1)
3. Sophie does not like to study alone for her exams. (Principle 2)
4. Sophie and Franck will take the same biology class next semester. (Principle 3)

(41)

### **IRF5déjeuner – Q5**

Conversation entre deux collègues de travail, Jean-Pierre (homme) et Christiane (femme). Ils discutent de là où ils pourraient manger ce midi.

Jean-Pierre: Christiane, est-ce que tu as décidé où tu voulais déjeuner?

Christiane: Je pensais aller au «Bon Pain» pour prendre un sandwich.

JP: Oh. J'avais pensé au nouveau restaurant avec une terrasse sur la plage.

Christiane: Je n'ai pas très envie d'être exposée au soleil.

(41)

1. Christiane ne veut pas aller au restaurant sur la plage. (CORRECT)
2. Christiane voudrait bien aller au nouveau restaurant sur la plage. (Principle 1)
3. Christiane veut bronzer pendant qu'elle déjeune avec Jean-Pierre. (Principle 2)
4. Christiane and Jean-Pierre savent où ils vont aller déjeuner. (Principle 3)

(37)

### **IRF5déjeuner**

This is a conversation between two co-workers, Jean-Pierre (male speaker) and Christiane (female speaker). They are talking about where they want to go for lunch.

Jean-Pierre: Christiane, have you decided where you want to have lunch today?

Christiane: I thought about going to the "Bon Pain" for a sandwich.

JP: Oh. I was thinking about the new restaurant with a patio right on the beach.

Christiane: I don't feel like being exposed to the sun.

(46)

1. Christiane doesn't want to go the restaurant by the beach. (CORRECT)
2. Christiane wants to have lunch at that new restaurant. (Principle 1)
3. Christiane wants to sunbathe while having lunch with Jean-Pierre. (Principle 2)
4. Christiane and Jean-Pierre know where they will have lunch. (Principle 3)

(37)

### **IRF6devoir – Q6**

Conversation entre une étudiante, Stéphanie (femme) et son professeur, M. Bodnar (homme). Stéphanie veut prendre rendez-vous avec M. Bodnar pour lui poser des questions sur un devoir à rendre.

Stéphanie: Bonjour M. Bodnar! J'ai une question au sujet du devoir à rendre, et j'aurais voulu prendre rendez-vous pour en discuter.

M. Bodnar: Oui bien sûr, quand êtes-vous libre?

Stéphanie: Est-ce que mardi matin vous irait?

M. Bodnar: Je ne vais pas à l'université le mardi.

(42)

1. M. Bodnar et Stéphanie ne vont pas se rencontrer mardi. (CORRECT)
2. M. Bodnar et Stéphanie sont tous les deux libres mardi. (Principle 1)

3. M. Bodnar enseigne à l'université tous les mardis. (Principe 2)
4. Stéphanie ne doute pas qu'elle pourra rendre son devoir au cours prochain. (Principe 3)

(40)

### **IRF6devoir**

This is a conversation between a student, Stéphanie (female speaker) and her advisor, M. Bodnar (male speaker). Stéphanie wants to meet Pierre to ask him questions about her homework.

Stéphanie: Hello, Mr. Bodnar! I have a question about the upcoming homework assignment, and I'd like to meet with you to discuss it.

M. Bodnar: Of course, when are you available?

Stéphanie: Would Tuesday morning work?

M. Bodnar: I don't come to university on Tuesdays.

(41)

1. Mr. Bodnar and Stéphanie are not going to meet on Tuesday. (CORRECT)
2. Mr. Bodnar and Stéphanie are both available on Tuesday. (Principe 1)
3. Mr. Bodnar teaches at university on Tuesdays. (Principe 2)
4. Stéphanie has no doubt that she can turn in her homework next class. (Principe 3)

(40)

### **IRF7supermarché - Q7**

Conversation entre deux colocataires, Charlie (homme) et Danielle (femme). Ils parlent d'aller faire les courses ensemble

Charlie: Danielle, nous n'avons plus beaucoup de choses à manger. Il faudrait aller au supermarché bientôt.

Danielle: Oui, tu as raison. Il n'y a pas beaucoup de nourriture dans le frigo.

Charlie: Est-ce que tu voudrais aller au supermarché cet après-midi?

Danielle: Le mardi je vais à mon cours de yoga.

(47)

1. Danielle n'ira pas au supermarché cet après-midi. (CORRECT)
2. Danielle et Charlie vont aller au supermarché cet après-midi. (Principe 1)
3. Danielle suit un cours de yoga tous les jours. (Principe 2)
4. Danielle et Charlie ont beaucoup de nourriture dans leur frigo. (Principe 3)

(35)

### **IRF7supermarché**

This is a conversation between two roommates, Charlie (male speaker) and Danielle (female speaker). They are talking about going grocery shopping together.

Charlie: Danielle, we are running out of things to eat. We need to go to the supermarket soon.

Danielle: Yes, you are right. There is not much food left in the fridge.

Charlie: Would you want to go to the supermarket this afternoon?

Danielle: On Tuesdays I go to my yoga class.



(48)

1. Danielle will not go grocery shopping this afternoon. (CORRECT)
2. Danielle and Charlie will go to the supermarket this afternoon. (Principle 1)
3. Danielle attends yoga classes every day. (Principle 2)
4. Danielle and Charlie have a lot of food in the fridge. (Principle 3)

(35)

### **IRF8exercice - Q8**

Conversation entre deux amis, Jacques (homme) et Sandrine (femme). Ils parlent de faire du sport ensemble

Jacques: Salut, Sandrine! Quelle belle journée pour faire du sport à l'extérieur aujourd'hui.

Sandrine: Oui, c'est le temps parfait pour faire du jogging au parc.

Jacques: Est-ce que tu veux faire du jogging avec moi cet après-midi?

Sandrine: J'ai déjà dit à une amie que j'allais jouer au tennis avec elle.

(47)

1. Sandrine ne peut pas faire du jogging avec Jacques aujourd'hui. (CORRECT)
2. Sandrine et Jacques feront du jogging ensemble au parc cet après-midi. (Principle 1)
3. Sandrine n'aime pas le tennis. (Principle 2)
4. Sandrine et Jacques vont jouer au basket cet après-midi. (Principle 3)

(35)

### **IRF8exercice**

This is a conversation between two friends, Jacques (male speaker) and Sandrine (female speaker). They are talking about exercising together.

Jacques: Hi, Sandrine! It is a beautiful day to work out outside today.

Sandrine: Yes, it's the perfect weather to go jogging in the park.

Jacques: Do you want to go jogging with me this afternoon?

Morena: I already told my friend I would play tennis with her.

(44)

1. Sandrine cannot go jogging with Jacques today. (CORRECT)
2. Sandrine and Jacques are going jogging in the park together this afternoon. (Principle 1)
3. Sandrine does not like tennis. (Principle 2)
4. Sandrine and Jacques are going to play basketball together this afternoon. (Principle 3)

(35)

### **IRF9buffet - Q9**

Conversation entre deux amis, Éric (homme) et Estelle (femme). Ils discutent de la possibilité d'aller au buffet à volonté de China King.

Éric: Quand est-ce qu'on va à China King pour essayer leur buffet à volonté?

Estelle: Quand tu veux Éric.

Éric: Demain à 19h, ça te va?

Estelle: Ben, j'ai des billets pour aller au concert de Drake à 19h15.

(35)

1. Estelle ne mangera pas à China King avec Éric demain à 19h. (CORRECT)
2. Estelle and Éric iront à China King ensemble demain à 19h. (Principle 1)
3. Estelle va à un concert cet après-midi. (Principle 2)
4. Estelle n'aime pas les buffets à volonté. (Principle 3)

(37)

### **IRF9buffet**

This is a conversation between two friends, Eric (male speaker) and Estelle (female speaker). They are talking about making plans to go to China King to try their all-you-can-eat buffet.

Eric: When are we going to China King to try their buffet?

Estelle: Whenever you want, Eric.

Eric: What about going tomorrow at 7PM?

Estelle: I have tickets to go to Drake's concert at 7:15pm.

(31)

5. Estelle won't try China King's buffet with Eric tomorrow at 7PM. (CORRECT)
6. Estelle and Eric will go to China King together tomorrow at 7PM. (Principle 1)
7. Estelle is going to a concert this afternoon. (Principle 2)
8. Estelle does not like all-you-can-eat buffets. (Principle 3)

(37)

### **IRF10pluie – Q10**

Conversation entre deux colocataires, Adrien (homme) et Mélanie (femme). Ils discutent de comment ils vont se rendre à l'université aujourd'hui

Adrien: Mélanie, on va à l'université à pied? Tu es prête?

Mélanie: Oui. Mais est-ce que tu as vu qu'il pleut beaucoup dehors.

Adrien: C'est pas un problème pour moi. J'ai l'habitude de marcher sous la pluie.

Mélanie: Je ne veux pas être mouillée.

(40)

1. Mélanie n'ira pas à l'université à pied aujourd'hui. (CORRECT)
2. Mélanie et Adrien iront à l'université à pied aujourd'hui. (Principle 1)
3. Ça ne dérange pas Mélanie d'être trempée en marchant à la fac. (Principle 2)
4. Mélanie et Adrien ne sont pas prêts à partir. (Principle 3)

(38)

### **IRF10pluie**

This is a conversation between two roommates, Adrien (male speaker) and Mélanie (female speaker). They are talking about how to go to campus today.

Adrien: Mélanie, are you ready to walk to campus?

Mélanie: Yes, I am. But did you see that it is raining pretty heavily.  
Adrien: That's not a problem for me. I am used to walking in the rain.  
Mélanie: I do not want to get soaked.  
(42)

1. Mélanie will not walk to campus today. (CORRECT)
  2. Mélanie and Adrien will go on foot to campus together today. (Principle 1)
  3. Mélanie doesn't mind getting soaked while walking to campus. (Principle 2)
  4. Mélanie and Adrien are not ready to go to campus yet. (Principle 3)
- (38)

### **IRF11dessert—Q 11 [used as practice 1]**

Conversation entre un couple, André (homme) et Carine (femme). Ils sont chez eux en train de diner et André offre des fraises à sa compagne

André: J'espère que tu as aimé le diner, mon amour. Est-ce que tu veux un dessert?  
Carine: Oui, j'ai adoré le diner. Merci. Et oui, je veux bien manger quelque chose pour le dessert.  
André: Parfait. J'ai des fraises ou du gâteau au chocolat.  
Carine: Ben tu sais bien que j'aime pas les fraises  
(50)

1. Carine voudrait une part de gâteau au chocolat pour le dessert. (CORRECT)
  2. Carine mangera des fraises après le diner. (Principle 1)
  3. Carine mange toujours des fraises après le diner. (Principle 2)
  4. Carine et André n'ont pas envie de manger de dessert. (Principle 3)
- (36)

### **IRF11dessert**

This is a conversation between a couple, André (male speaker) and Carine (female speaker). They are having dinner at home and André offers strawberries to his girlfriend.

André: I hope you liked the dinner, my love. Would you like some dessert?  
Carine: Yes, I loved the dinner. Thank you. And yes, I'd like to have something for dessert.  
André: Perfect. I have strawberries or chocolate cake.  
Carine: Well, you know that I don't like strawberries.  
(44)

1. Carine would like to have a slice of chocolate cake after dinner. (CORRECT)
  2. Carine will eat some strawberries after dinner. (Principle 1)
  3. Carine always eats strawberries after dinner. (Principle 2)
  4. Carine and André do not feel like eating dessert after dinner. (Principle 3)
- (36)

### **IRF12notes - Q 12**

Conversation entre deux camarades de classe, Frédéric (homme) et Anna (femme). Frédéric a manqué des cours et demande à Anna s'il peut avoir ses notes de classe.

Frédéric: Salut Anna, j'étais absent en classe la semaine dernière.

Anna: Ouais, j'ai vu que t'étais pas là. Tu étais malade?

Frédéric: Non en fait j'étais en vacances. Est-ce que tu pourrais me passer tes notes de cours?

Anna: C'est que je n'ai pas l'habitude de prêter mes notes.

(45)

1. Anna ne va pas prêter ses notes à Frédéric. (CORRECT)
2. Anna veut bien que Frédéric utilise ses notes. (Principe 1)
3. Anna aime partager ses notes avec ses camarades. (Principe 2)
4. Anna et Frédéric sont allés en vacances ensemble la semaine dernière. (Principe 3)

(36)

### **IRF12notes**

This is a conversation between two classmates, Frédéric (male speaker) and Anna (female speaker). Frédéric missed class, and he is asking Anna for her notes.

Frédéric: Anna, I couldn't come to class last week.

Anna: Yeah, I noticed you weren't there. Were you sick?

Frédéric: Actually, I was on vacation. Would you mind lending me your notes from last week?

Anna: Well, I'm not used to sharing my notes.

(41)

1. Anna will not lend her class notes to Frédéric. (CORRECT)
2. Anna does not mind Frédéric using her class notes. (Principe 1)
3. Anna always likes sharing her notes with her classmates. (Principe 2)
4. Anna and Frédéric went on vacation together last week. (Principe 3)

(36)

### **IRF13chaussures – Q 13**

Conversation entre deux amis, Daniel (homme) et Béatrice (femme). Béatrice a besoin d'une nouvelle paire de chaussure pour un mariage.

Béatrice: Daniel, j'ai besoin d'acheter une nouvelle paire de chaussures pour le mariage de Sandrine.

Daniel: Mais est-ce que tu ne t'es pas déjà acheté une nouvelle paire la semaine dernière?

Béatrice: Si, mais j'en ai besoin d'une autre. Tu veux venir avec moi pour l'acheter.

Daniel: Tu ne devrais pas dépenser encore plus d'argent pour des chaussures.

(54)

1. Daniel ne veut pas accompagner Béatrice pour qu'elle s'achète de nouvelles chaussures. (CORRECT)
2. Daniel va aider Béatrice à s'acheter des nouvelles chaussures. (Principe 1)
3. Daniel pense que Béatrice devrait dépenser plus d'argent pour des chaussures. (Principe 2)
4. Daniel and Béatrice vont chacun s'acheter des nouvelles chaussures. (Principe 3)

300

(41)

### **IRF13chaussures**

This is a conversation between two friends, Daniel (male speaker) and Béatrice (female speaker). Béatrice needs a new pair of shoes for a wedding.

Béatrice: Daniel, I need to buy a new pair of shoes for Sandrine's wedding.

Daniel: But didn't you buy yourself a new pair last week?

Béatrice: Yes, but I need a new pair anyways. Do you want to come with me to buy them?

Daniel: You shouldn't spend more money on shoes.

(48)

1. Daniel won't go with Béatrice to buy a new pair of shoes. (CORRECT)
2. Daniel will help Béatrice find a new pair of shoes. (Principle 1)
3. Daniel thinks Béatrice should spend money on shoes. (Principle 2)
4. Daniel and Béatrice will buy a new pair of shoes each. (Principle 3)

(41)

### **IRF14snacks - Q14**

Conversation entre deux colocataires, Damien (homme) et Isabelle (femme). Ils sont au cinéma et vont commander quelque chose à manger avant que le film ne commence.

Damien: Isabelle, est-ce que tu voudrais acheter quelque chose à manger avant le film?

Isabelle: Oui, c'est une très bonne idée parce que j'ai un peu faim!

Damien: Je pensais acheter un grand sac de popcorn pour nous deux.

Isabelle: Je préférerais manger une glace.

(41)

1. Isabelle ne veut pas acheter de popcorn avant le film. (CORRECT)
2. Isabelle veut bien partager un sac de popcorn avec Damien. (Principle 1)
3. Isabelle n'aime pas la glace. (Principle 2)
4. Ni Isabelle ni Damien ne veut acheter quelque chose à manger avant le film. (Principle 3)

(39)

### **IRF14snacks**

This is a conversation between two roommates, Damien (male speaker) and Isabelle (female speaker). They are at the movies and they are going to order something to eat before the movie starts.

Damien: Isabelle, would you like to buy something to eat before the movie starts?

Isabelle: Yes. I think it is a great idea because I am kind of hungry!

Damien: I was thinking about getting a big bag of popcorn to share.

Isabelle: I would rather have ice cream.

(45)

1. Isabelle doesn't want to buy popcorn before the movie starts. (CORRECT)
2. Isabelle will share a bag of popcorn with Damien. (Principle 1)

301

3. Isabelle does not like ice cream. (Principle 2)
4. Neither Isabelle nor Damien want to buy something to eat while at the movies. (Principle 3)

(39)

### **IRF15vacances – Q15**

Conversations entre deux étudiants, Jérémy (homme) and Erika (femme). Ils discutent de destinations de vacances.

Jérémy: Salut, Erika! Alors, tu penses partir en vacances bientôt?

Erika: Oui! J'ai vraiment envie de faire un voyage! Tu as des suggestions?

Jérémy: Je pense que tu devrais aller à New York City. Tu vas adorer!

Erika: J'ai déjà été à New York.

(40)

1. Erika n'ira pas à New York City pendant les vacances. (CORRECT)
2. Erika ira à New York City pendant les vacances. (Principle 1)
3. Erika n'a jamais été à New York City, mais elle aimerait bien y aller. (Principle 2)
4. Erika et Jérémy veulent faire un voyage ensemble. (Principle 3)

(41)

### **IRF15vacances**

This is a conversation between two students, Jérémy (male speaker) and Erika (female speaker). They are talking about destinations they could go on vacation.

Jérémy: Hi, Erika! So, you are thinking about going on vacations soon?

Erika: Yes! I really want to go on a trip! Do you have any suggestions?

Jérémy: I think you should go to New York City. You will love it!

Erika: I've already been to New York City.

(45)

1. Erika won't go to New York City on vacation. (CORRECT)
2. Erika will go to New York City on vacation. (Principle 1)
3. Erika has never been to New York City, but she'd like to go. (Principle 2)
4. Erika and Jérémy want to go on a trip together. (Principle 3)

(41)

### **IRF16musée – Q16**

Conversation entre deux bons amis, Maurice (homme) et Marianne (femme). Ils parlent de visiter un musée

Maurice: Marianne, il fait mauvais aujourd'hui, mais je voudrais quand même sortir.

Marianne: Oui, c'est vrai. Quel temps horrible! Mais moi aussi j'aimerais aller quelque part.

Maurice: Est-ce que veux aller au Musée des Sciences Naturelles?

Marianne: Je préfère aller au Musée d'Art Moderne.

(40)

1. Marianne n'a pas vraiment envie d'aller au Musée des Sciences Naturelles aujourd'hui. (CORRECT)
2. Marianne voudrait bien aller visiter le Musée des Sciences Naturelles. (Principe 1)
3. Marianne n'aime pas le Musée d'Art Moderne. (Principe 2)
4. Marianne et Maurice veulent aller dans un parc d'attractions ensemble aujourd'hui. (Principe 3)

(40)

### **IRF16musée**

This is a conversation between two old friends, Maurice (male speaker) and Marianne (female speaker). They are talking about visiting a museum.

Maurice: Marianne, it's awful today, but I still want to go out.

Marianne: Yes, that's true. It's horrible today, but I'd also like to go somewhere.

Maurice: Do you want to go to the Museum of Natural Sciences?

Marianne: I'd rather go to the Museum of Modern Art.

(44)

1. Marianne does not feel like going to the Museum of Natural Sciences. (CORRECT)
2. Marianne wants to visit the Museum of Natural Sciences. (Principe 1)
3. Marianne does not like the Museum of Modern Art. (Principe 2)
4. Marianne and Maurice want to go to an amusement park. (Principe 3)

(40)

### **Irony (IRO)**

#### **IRO1volume - Q17**

Conversation entre un père et sa fille. Ils regardent la télé à la maison

Père: Je vais me coucher. Je suis fatigué.

Fille: Tu ne veux pas regarder le film avec moi?

Père: J'ai sommeil et je dois me lever très tôt demain matin.

*(Quelques secondes plus tard, le père crie de sa chambre):*

Père: Le volume est tellement bas que je ne peux pas entendre depuis ma chambre!

(41)

1. Le père pense que la télé est trop forte. (CORRECT)
2. Le père n'entend pas bien la télé. (Principe 1)
3. Le père pense que la télé est réglée au bon volume. (Principe 2)
4. Le père pense qu'il devrait regarder un film avec sa fille. (Principe 3)

(38)

#### **IRO1volume (17)**

This is a conversation between a father and his daughter. They are at home watching TV at night.

Father: I am going to bed. I am tired.

Daughter: Don't you want to watch the movie with me?  
Father: I'm sleepy and I have to get up really early tomorrow.  
*(A few seconds later, the father screams from his room)*  
Father: The volume is so low that I can't hear from my bedroom!  
(41)

1. The father thinks the TV's volume is quite loud. (CORRECT)
  2. The father cannot hear the TV very well. (Principle 1)
  3. The father thinks the TV is at the right volume. (Principle 2)
  4. The father thinks he should watch the movie with his daughter. (Principle 3)
- (38)

### **IR02climat Q18**

Conversation entre deux colocataires, Valérie (femme) et Marc (homme), le matin. Ils discutent du temps qu'il fait.

Valérie: Est-ce que tu es prêt? Vite. Dépêche-toi. J'ai cours dans 40 minutes.  
Marc: Oui, oui je suis prêt. Allons-y. Toi, tu es prête?  
Valérie: Oui, j'ai toutes mes affaires.  
Marc: Tu ne vas pas prendre ton manteau? Il fait si chaud dehors!  
(39)

1. Marc pense qu'il fait vraiment froid ce matin. (CORRECT)
  2. Marc pense vraiment que la matinée va être très chaude. (Principle 1)
  3. Marc pense que Valérie ne devrait pas amener son manteau en cours. (Principle 2)
  4. Marc pense qu'il pleut à verse ce matin. (Principle 3)
- (38)

### **IR02climat (18)**

This is a conversation between two roommates, Valérie (female speaker) and Marc (male speaker), in the morning. They are talking about the weather.

Valérie: Are you ready to go? Hurry up. I have class in 40 minutes.  
Marc: Yes, yes. I'm ready. Let's go. Are you ready?  
Valérie: Yes. I already got all my things.  
Marc: Aren't you taking your coat with you? It's so hot outside!  
(40)

1. Marc thinks that it's really cold this morning. (CORRECT)
  2. Marc thinks that they're having quite a hot morning. (Principle 1)
  3. Marc thinks that Valérie should not take her coat to go to class. (Principle 2)
  4. Marc thinks that it's raining hard this morning. (Principle 3)
- (38)

### **IR03devoirs – Q19**

Conversation entre deux colocataires, Ethan (homme) et Sandra (femme). Ils parlent de la quantité de travail qu'ils ont à faire.



Ethan: Comment s'est passé ta journée? Tu avais une présentation aujourd'hui, non?  
Sandra: Oui; j'ai eu une journée super longue. Mais ma présentation s'est bien passée.  
Et toi?  
Ethan: Génial! Malheureusement pour moi, j'ai beaucoup de devoirs donc je vais devoir  
me coucher tard ce soir!  
Sandra: Ah oui, c'est tellement rare!  
(48)

1. Ethan se couche souvent tard le soir. (CORRECT)
2. Ethan se couche toujours tôt le soir. (Principe 1)
3. Il est rare que Sandra ait des présentations à faire. (Principe 2)
4. Ethan et Sandra ont tous les deux fait une présentation en classe aujourd'hui.  
(Principe 3)  
(37)

### IR03devoirs (19)

This is a conversation between two roommates, Ethan (male speaker) and Sandra (female speaker). They are talking about the amount of work they have to do.

Ethan: How was your day? You said you had a presentation today, right?  
Sandra: Yes. My day was really long. At least the presentation went well. How about  
you?  
Ethan: That's great! Unfortunately, I have a lot of homework so I'll have to stay up late.  
Sandra: Oh, that's so rare!  
(47)

1. Ethan is used to going to bed very late. (CORRECT)
2. Ethan always goes to bed very early. (Principe 1)
3. It is quite unusual for Sandra to have presentations in class. (Principe 2)
4. Ethan and Sandra both had a presentation in class today. (Principe 3)  
(37)

### IR04télévision - Q20

Conversation entre deux colocataires, Abdel (homme) et Elsa (femme). Ils discutent de ce  
qu'ils pourraient regarder à la télé ensemble

Abdel: Elsa, qu'est-ce que tu vas faire ce soir?  
Elsa: Rien, pourquoi?  
Abdel: Il y a une nouvelle série à la télé que j'aimerais bien regarder. Tu veux la  
regarder avec moi?  
Elsa: Tu me connais, à chaque fois que quelqu'un allume la télé, je vais dans ma  
chambre pour lire. Je trouve que la télé est super éducative.  
(55)

1. Elsa pense que la télé n'est pas éducative. (CORRECT)
2. Elsa pense qu'on peut apprendre beaucoup avec la télé. (Principe 1)
3. Elsa regarde des émissions éducatives à la télé tous les soirs. (Principe 2)
4. Elsa et Abdel vont regarder une nouvelle série ensemble ce soir. (Principe 3)  
(39)

### IR04télévision (20)

This is a conversation between two roommates, Abdel (male speaker) and Elsa (female speaker). They are talking about watching a TV show together.

Abdel: Elsa, what are you going to do tonight?

Elsa: Nothing, why?

Abdel: There's a new TV show that I'd like to watch. Do you want to watch it with me?

Elsa: You know me, every time someone switches the TV on, I go to my room and read. I find TV really educational.

(50)

1. Elsa thinks that TV is not educational. (CORRECT)
2. Elsa thinks people can learn a lot from watching TV. (Principle 1)
3. Elsa watches educational programs on TV every single night. (Principle 2)
4. Elsa and Abdel are going to watch a new TV show together tonight. (Principle 3)

(39)

### IR05Europe - Q21

Conversation entre deux amis, Richard (homme) et Alicia (femme). Ils discutent des projets de voyage de Richard

Richard: Alicia, est-ce que je t'ai dit que je vais partir en voyage le mois prochain?

Alicia: Non, je ne savais pas que tu allais partir en vacances! Où est-ce que tu vas aller?

Richard: Plusieurs villes en Europe: Amsterdam, Barcelone, Milan, Paris, Athènes, Berlin et Prague.

Alicia: Oh non, arrête, tu as la vie dure!

(52)

1. Alicia pense que Richard a la belle ville. (CORRECT)
2. Alicia pense que Richard est bien malheureux. (Principle 1)
3. Alicia pense que Richard a vécu des choses dures dans sa vie. (Principle 2)
4. Alicia and Richard projettent de voyager ensemble en Europe. (Principle 3)

(36)

### IR05Europe (21)

This is a conversation between two friends, Richard (male speaker) and Alicia (female speaker). They are talking about the trip Ricardo is planning to make.

Richard: Alicia, did I tell you that I'm going on a trip next month?

Alicia: No! I didn't know you were going on vacation! Where are you going to go?

Richard: Some cities in Europe: Amsterdam, Barcelona, Milan, Paris, Athens, Berlin and Prague.

Alicia: Oh please, stop, you have it so hard.

(48)

1. Alicia thinks Richard's life is pretty good. (CORRECT)
2. Alicia thinks Richard is quite unhappy. (Principle 1)
3. Alicia thinks Richard went through some really hard things in his life. (Principle 2)

4. Alicia and Richard plan to make a trip to Europe together. (Principle 3)  
(36)

### IR06diner—Q22

Conversation entre deux amis, Olivier (homme) et Marie (femme), au sujet du repas que Marie a préparé

Marie: Olivier, tu veux du dessert? J'ai fait une tarte aux fraises.

Olivier: Ah? Oui bien sûr. J'adore la tarte aux fraises.

Marie: Et qu'est-ce que tu as pensé du poulet-frites que j'ai préparé?

Olivier: Ben en fait, ton repas était vraiment horrible. J'ai presque rien mangé!

(43)

1. Olivier a beaucoup mangé pour le dîner. (CORRECT)
2. Olivier n'a pas apprécié le poulet et les frites préparés par Marie. (Principle 1)
3. Olivier pense que Marie est une cuisinière horrible. (Principle 2)
4. Olivier et Marie ont commandé une tarte aux fraises au restaurant. (Principle 3)  
(38)

### IR06diner (22)

This is a conversation between two friends, Olivier (male speaker) and Maria (female speaker). They are talking about the food Marie prepared for dinner.

Marie: Olivier, do you want some dessert? I made a strawberry tart.

Olivier: Why? Of course. I love strawberry tart.

Marie: And, how did you like the chicken and fries that I baked?

Olivier: Actually, your food was horrible. I barely ate!

(38)

1. Olivier ate a lot for dinner. (CORRECT)
2. Olivier did not like the chicken with fries that Marie baked for dinner. (Principle 1)
3. Olivier thinks Marie is a horrible cook. (Principle 2)
4. Olivier and Marie ordered a strawberry cake for dessert at a restaurant. (Principle 3)  
(38)

### IR07chaleur - Q23

Conversation entre deux amis, Stéphane (homme) et Margot (femme). Ils discutent du temps qu'il fait

Margot: Stéphane, tu es prêt pour aller au musée? Il fait soleil donc je pense qu'on peut y aller à pied.

Stéphane: Ah oui, c'est une bonne idée de marcher. Allons-y.

*(Une fois à l'extérieur)*

Margot: Oh la la, qu'est-ce qu'il fait chaud aujourd'hui!

Stéphane: Heureusement que tu me le dis, je n'aurais pas remarqué!

(47)

1. Stéphane avait remarqué qu'il faisait chaud aujourd'hui. (CORRECT)

2. Stéphane ne savait pas qu'il faisait chaud aujourd'hui. (Principe 1)
3. Stéphane s'est rendu compte qu'ils allaient se rendre au musée à pied. (Principe 2)
4. Margot et Stéphane vont prendre le bus pour aller au musée parce qu'il pleut. (Principe 3)  
(41)

### IR07chaleur (23)

This is a conversation between two friends, Stéphane (male speaker) and Margot (female speaker). They are talking about the weather.

Margot: Stéphane, are you ready to go to the museum? It's sunny, so I think we can walk there.

Stéphane: I think walking there is a great idea. Let's go.

*(Once out)*

Margot: Oh my, it's so hot today!

Stéphane: I'm so glad you told me, I would have never noticed!

(45)

1. Stéphane was aware that it was hot today. (CORRECT)
2. Stéphane did not know that it was hot today. (Principe 1)
3. Stéphane realized they were going to walk to the museum. (Principe 2)
4. Margot and Stéphane will go to the museum by bus because it is raining. (Principe 3)  
(41)

### IR08grandesécoles - Q24

Conversation entre deux amis, Paul (homme) et Iris (femme). Ils discutent de leur demande d'inscription à des grandes écoles

Paul: Salut, Iris! Est-ce que tu as déjà envoyé des demandes d'inscription aux grandes écoles?

Iris Oui je les ai envoyées la semaine dernière! Et toi? Des nouvelles?

Paul: Non, mais je suis sûr que je vais être accepté partout. Je suis un excellent étudiant.

Iris: Oui, c'est ça, et moi je suis Harry Potter.

(51)

1. Iris ne pense pas que Paul soit un excellent étudiant. (CORRECT)
2. Iris pense que Paul est un étudiant excellent. (Principe 1)
3. Iris et Paul vont regarder Harry Potter ensemble. (Principe 2)
4. Iris et Paul savent déjà dans quelles écoles ils vont poursuivre leurs études. (Principe 3)  
(39)

### IR08grandesécoles (24)

This is a conversation between two friends, Paul (male speaker) and Iris (female speaker). They are talking about their college applications.

Paul: Hi, Iris! Have you already sent your applications to the universities?

Iris Yes, I submitted them last week! What about you? Any news?

Paul: No, but I'm sure I'll be accepted by all of them. I'm an excellent student.

Iris: yes, that's right and I'm Harry Potter.

(44)

1. Iris does not think Paul is an excellent student. (CORRECT)
  2. Iris believes Paul is an excellent student. (Principle 1)
  3. Iris and Paul will go to watch Harry Potter together. (Principle 2)
  4. Iris and Paul already know in which universities they will complete their degrees. (Principle 3)
- (39)

### **IR09corrections - Q25**

Conversation entre deux collègues, Lucas (homme) et Mélanie (femme). Ils parlent de la quantité de travail qu'ils doivent faire pendant le week-end

Lucas: Mélanie, tu es libre ce week-end? Il y a un nouveau film que j'aimerais bien voir.

Mélanie: Je ne pense pas que je puisse y aller. J'ai trop de travail.

Lucas: Tu dois encore corriger tous les essais de tes étudiants?

Mélanie: Oui! je suis trop contente de devoir corriger 50 copies!

(49)

1. Mélanie n'aime pas avoir à corriger 50 copies de ses étudiants. (CORRECT)
  2. Ça ne dérange pas Mélanie d'avoir à corriger 50 copies de ses étudiants. (Principle 1)
  3. Mélanie doit encore écrire 50 essais. (Principle 2)
  4. Lucas veut aider Mélanie à corriger ses essais pendant le week-end. (Principle 3)
- (41)

### **IR09corrections (25)**

This is a conversation between two colleagues, Lucas (male speaker) and Mélanie (female speaker). They are talking about the amount of work they have to do during the weekend.

Lucas: Mélanie, are you busy this weekend? There's a new movie out that I want to watch.

Mélanie: I don't think I can go. I have too much work.

Lucas: Do you still have to grade your students' essays?

Mélanie: Yes! Having 50 papers to grade makes me so happy!

(46)

1. Mélanie does not like having 50 essays from her students to grade. (CORRECT)
  2. Mélanie does not mind having 50 essays from her students to grade. (Principle 1)
  3. Mélanie has to write 50 essays. (Principle 2)
  4. Lucas wants to help Mélanie to grade essays during the weekend. (Principle 3)
- (41)

### **IR010entorse - Q26**

Conversation entre deux colocataires, Emmanuel (homme) et Camille (femme). Ils parlent du fait qu'Emmanuel se soit foulé la cheville en faisant du jogging

Camille: Est-ce que tu as fini tes cours pour aujourd'hui? Qu'est-ce que tu vas faire après?

Emmanuel: Oui j'ai fini. Je vais aller faire du jogging maintenant.

*(Une heure plus tard - quand Emmanuel rentre à la maison)*

Camille: Oh la la, Emmanuel! Est-ce que tu as vu ton pied? Il est tout gonflé!

Emmanuel: Ah ouais sérieux? Je n'avais pas vu.

(47)

1. Emmanuel avait remarqué que son pied était enflé. (CORRECT)
2. Emmanuel ne savait pas que son pied était gonflé. (Principe 1)
3. Camille n'a pas remarqué que le pied d'Emmanuel était enflé. (Principe 2)
4. Emmanuel et Camille ont fait du jogging ensemble. (Principe 3)

(35)

### IR010entorse (26)

This is a conversation between two roommates, Emmanuel (male speaker) and Camille (female speaker). They are talking about Emmanuel having sprained his ankle while jogging.

Camille: Have you finished teaching for today? What are you going to do next?

Emmanuel: Yes, I'm done. I am going to for a jog now.

*(One hour later - When Emmanuel comes back home)*

Camille: Oh my god, Emmanuel! Did you see your foot? It's so swollen!

Emmanuel: Seriously? I hadn't noticed.

(40)

1. Manuel noticed that his foot was very swollen. (CORRECT)
2. Manuel did not know that his foot was very swollen. (Principe 1)
3. Camille did not notice that Manuel's foot was very swollen. (Principe 2)
4. Manuel and Camille went jogging together. (Principe 3)

(34)

### IR011ponctualité - Q27

Conversation entre deux étudiants, Nicolas (homme) et Claudia (femme). Claudia téléphone à Nicolas pour lui dire qu'elle sera en retard à leur rendez-vous au café.

Nicolas: Salut, Claudia. Est-ce que tu vas arriver dans longtemps? Je suis déjà au café et j'ai réservé une table?

Claudia: Nicolas, je suis presque arrivée. Le bus était en retard mais je ne suis pas très loin.

*(Claudia entre dans le café et s'approche de la table où Nicolas est assis)*

Claudia: Je suis là.

Nicolas: Toujours à l'heure.

(42)

1. Claudia arrive souvent en retard à ses rendez-vous. (CORRECT)

2. Claudia arrive toujours à l'heure à ses rendez-vous. (Principe 1)
  3. Le bus de Claudia est arrivé à l'heure. (Principe 2)
  4. Claudia et Nicolas ont pris le bus ensemble pour aller au café. (Principe 3)
- (36)

### IR011ponctualité (27)

This is a conversation between two classmates, Nicolas (male speaker) and Claudia (female speaker). Claudia calls Nicolas to let him know she will be late for their meeting at a café.

Nicolas: Hi, Claudia. Will it take you much longer to arrive? I am already at the café and I got a table.

Claudia: Nicolas, I am almost there. The bus came a couple of minutes late, but I'm not far now.

*(Claudia comes into the café and comes closer to the table where Nicolas is)*

Claudia: I'm here.

Nicolas: Always on time!

(45)

1. Claudia tends to arrive late to her meetings. (CORRECT)
  2. Claudia always arrives on time to her meetings. (Principe 1)
  3. The bus that Claudia took was on time. (Principe 2)
  4. Claudia and Nicolas took the bus together to get to the café. (Principe 3)
- (36)

### IR012matin - Q28

Discussion entre couple, Maxime (homme) et Elisabeth (femme). Ils discutent de leur nuit au réveil

Maxime: Bonjour, chéri. Est-ce que tu as fait des cauchemars cette nuit? Tu as beaucoup bougé et parlé à plusieurs reprises. Moi j'ai bien dormi.

Elisabeth: Bonjour mon cœur. Je ne me rappelle pas mais j'ai mal à la tête maintenant.

Maxime Alors est-ce que tu veux un café?

Elisabeth: Regarde-moi: je suis complètement réveillée là!

(52)

1. Elisabeth a besoin d'un café parce qu'elle a mal dormi. (CORRECT)
  2. Elisabeth n'a pas besoin de café car elle a très bien dormi. (Principe 1)
  3. Elisabeth est toujours bien réveillée le matin. (Principe 2)
  4. Elisabeth et Maxime ont tous les deux eu des cauchemars cette nuit. (Principe 3)
- (41)

### IR012matin (28)

This is a conversation between a couple, Maxime (male speaker) and Elisabeth (female speaker). They are talking when they wake up in the morning.

Maxime: Good morning, babe. Did you have nightmares last night? You were restless and you talked at times. I slept well.

Elisabeth: Good morning, sweetheart. I do not remember but I have a headache.

Maxime So, do I make some coffee?  
Elisabeth: Look at me: I am so wide awake!  
(46)

1. Elisabeth would like to have a coffee in the morning because she did not sleep well. (CORRECT)
2. Elisabeth does not need coffee now because she slept well. (Principle 1)
3. Elisabeth is wide awake in the morning. (Principle 2)
4. Elisabeth and Maxime both had nightmares last night. (Principle 3)  
(41)

### IR013fête – Q29

Conversation entre deux amis, Carl (homme) et Mathilde (femme). Ils parlent d'une fille qu'ils ont vu à la fête d'Anna.

Mathilde: Salut, Carl. Est-ce que tu sais à quelle heure la fête commence? Je veux vraiment y aller.

Carl: Salut, Mathilde. Oui, c'est à 21h chez Anna. Est-ce que tu veux y aller avec moi?

Mathilde: Oui s'il te plait, parce que j'ai un problème avec ma voiture.

*(Une fois arrive à la fête)*

Carl: Est-ce que tu as vu Caroline? Son visage est tellement naturel!

(56)

1. Carl pense que Caroline est trop maquillée. (CORRECT)
2. Carl pense que Caroline est peu maquillée. (Principle 1)
3. Carl croit qu'il y a un singe à la fête. (Principle 2)
4. Carl et Mathilde ont été à la fête de Natacha à 21h ce soir-là. (Principle 3)  
(38)

### IR013fête (29)

This is a conversation between two friends, Carl (male speaker) and Mathilde (female speaker). They are talking about a girl they have seen at Anna's party.

Mathilde: Hi, Carl. Do you know what time the party is? I really want to go.

Carl: Hi, Mathilde. Yes. It is at 9PM at Anna's place. Do you want to go with me?

Mathilde: Yes, please, I'm having some car issues.

*(As soon as they get to the party)*

Carl: Did you see Caroline? Her face looks really natural!

(48)

1. Carl thinks Caroline put on a lot of make-up. (CORRECT)
2. Carl thinks Caroline barely put any make up on. (Principle 1)
3. Carl believes there is a monkey at the party. (Principle 2)
4. Carl and Mathilde went to Natacha's party at 9PM that night. (Principle 3)  
(38)

### IR014froid – Q30



Conversation entre deux étudiants, Jean-Charles (homme) et Anabelle (femme). Ils discutent du fait qu'il fait froid dans leur salle de classe.

Jean-Charles: Anabelle, c'est génial le printemps est arrivé! J'en avais vraiment besoin!

Anabelle: Oui, Jean-Charles. Mais je dois quand même mettre une veste en classe.

Jean Charles: Ah bon pourquoi? Il fait bon dans la salle de cours!

Anabelle: Ah oui bien sûr. On a l'impression d'être dans un sauna!

(46)

1. Anabelle pense qu'il fait froid dans la salle de classe. (CORRECT)
2. Anabelle trouve qu'il fait chaud dans la salle de classe. (Principe 1)
3. Anabelle pense qu'il y a un sauna à l'université. (Principe 2)
4. Anabelle et Jean-Charles ne veulent pas que le printemps arrive. (Principe 3)

(39)

### IR014froid (30)

This is a conversation between two classmates, Jean-Charles (male speaker) and Anabelle (female speaker). They are talking about how cold it is in class.

Jean-Charles: Anabelle, it is great spring has arrived! I really need it!

Anabelle: Yes, Jean-Charles. But I still need to wear a jacket in class.

Jean-Charles: But why? It is really nice in class!

Anabelle: For sure. It is like being in a sauna!

(40)

1. Anabelle thinks that it is still cold in class. (CORRECT)
2. Anabelle feels warm when she is in class. (Principe 1)
3. Anabelle thinks that there is a sauna at the university. (Principe 2)
4. Anabelle and Jean-Charles do not want the spring season to arrive yet. (Principe

3)

(39)

### IR015chaise – Q31

Conversation entre deux colocataires, Antoine (homme) et Clara (femme). Ils discutent des meubles qu'Antoine a achetés hier dans un magasin d'occasions.

Antoine: J'ai acheté des meubles au magasin d'occasions près de chez nous hier. Qu'est-ce que tu en penses?

Clara: Super! Nous avons maintenant les meubles que TU voulais.

Antoine: Oui! On avait vraiment besoin de remplacer nos chaises. Essaie cette chaise. Qu'est-ce que tu en penses?

Clara: (*la chaise chancelle*). Elle n'a pas l'air vieille du tout!

(51)

1. Clara pense que la chaise n'est pas neuve car elle vacille. (CORRECT)
2. Clara pense que la chaise est neuve car elle est en bon état. (Principe 1)
3. La chaise n'est pas vieille du tout. (Principe 2)
4. Clara and Antoine ont acheté des meubles ensemble. (Principe 3)

(39)

### IR015chaise (31)

This is a conversation between two roommates, Antoine (male speaker) and Clara (female speaker). They are talking about the new furniture Antoine bought yesterday.

Antoine: I bought new furniture at the store on the corner yesterday. What do you think?

Clara: Fantastic! We now got the new furniture you wanted.

Antoine: Yes! The chairs needed to be changed urgently. Try this one. ¿How is it?

Clara: (*the chair wobbles*). It does not seem old at all!

(45)

1. Clara thinks the chair is not new because it wobbles. (CORRECT)
  2. Clara believes the chair is new because it is in good condition. (Principle 1)
  3. Clara believes the chair is not old at all. (Principle 2)
  4. Clara and Antoine bought the furniture together. (Principle 3)
- (38)

### IR016capuccino - Q32

Conversation entre deux amis, Joshua (homme) et Irène (femme). Ils discutent du cappuccino qu'Irène a commandé au café

Joshua: Salut, Irène! Je suis content que tu sois venue; ça faisait longtemps que je ne t'avais pas vu.

Irène: Salut, Joshua! Oui. Ça fait longtemps.

Joshua: J'ai déjà commandé. Qu'est-ce que tu voudrais boire?

Irène: Je prendrai un cappuccino avec de la crème Chantilly et de la sauce au chocolat.

Joshua: Ah ben je vois que tu n'aimes pas les boissons sucrées toi, non?!

(61)

1. Irène a commandé un cappuccino très sucré. (CORRECT)
  2. Irène a commandé un cappuccino qui n'est pas très sucré. (Principle 1)
  3. Irène a commandé une tasse de café noir. (Principle 2)
  4. Irène et Joshua aiment le chocolat. (Principle 3)
- (31)

### IR016capuccino (32)

This is a conversation between two friends, Joshua (male speaker) and Irène (female speaker). They are talking about the cappuccino Irene ordered at the café.

Joshua: Hi, Irène! I'm so happy you came: It's been so long since I've seen you.

Irène: Hi, Joshua! Yes. It's been a while.

Joshua: I already ordered. What would you like to drink?

Irène: I'll have a cappuccino with some whipped cream and chocolate syrup.

Joshua: Well sounds like you really don't like sugary drinks, now do you?

(52)

1. Irène ordered a very sweet cup of cappuccino. (CORRECT)
2. Irène ordered a cappuccino which was not sweet at all. (Principle 1)
3. Irène ordered a cup of black coffee. (Principle 2)

4. Irène and Joshua like chocolate. (Principle 3)  
(30)

## 2. Non-Conventional Implicatures: Indirect Opinions

### *Indirect opinions (OPI)*

#### **OPI1pizza – Q33**

Conversations entre deux collègues de travail, Stéphanie (femme) et Matthieu (homme) qui discutent de la possibilité de commander une pizza pour le déjeuner.

Stéphanie: Matthieu, on va commander une pizza. Tu veux participer?

Matthieu: Vous allez commander dans quel restaurant?

Stéphanie: A la «Tour de Pise», la pizzeria du quartier. Les pizzas sont bonnes et pas trop chères.

Matthieu: La dernière fois que j'en ai mangé une j'ai été malade.

(43)

- I. Matthieu pense que la pizzeria n'est pas bonne. (CORRECT)
  - II. Matthieu aime manger des pizzas à La Tour de Pise. (Principle 1)
  - III. Matthieu a dépensé peu d'argent la dernière fois qu'il a mangé à la Tour de Pise. (Principle 2)
  - IV. Matthieu pense que la pizzeria est loin. (Principle 3)
- (41)

#### **OPI1pizza (33)**

This is a conversation between two male co-workers, Stéphanie (female speaker) and Matthieu (male speaker). They are talking about ordering pizza for lunch.

Stéphanie: Matthieu, we're gonna order a pizza. Do you want to join us?

Matthieu: What place are you ordering it from?

Stéphanie: From the Tower of Pisa, the neighborhood pizza place. Their pizzas are pretty good and cheap.

Matthieu: Last time I had one, it made me sick.

(44)

1. Matthieu thinks the pizza place is not good. (CORRECT)
  2. Matthieu likes eating pizzas at the Tower of Pisa restaurant. (Principle 1)
  3. Matthieu paid very little the last time he ate at The Tower of Pisa. (Principle 2)
  4. Matthieu thinks the pizza place is too far. (Principle 3)
- (40)

#### **OPI2mariage – Q34**

Conversation entre deux amis, Rémy (homme) et Anna (femme). Ils discutent du mariage auquel Anna a assisté.

Rémy: Comment était la cérémonie de mariage de Lucie le week-end dernier?

Anna: C'était sympa. On s'est bien amusé!

Rémy: Et tu as aimé sa robe?

Anna: Si on était dans les années 80.

(30)

1. Anna n'aime pas la robe de Lucie (CORRECT)
  2. Anna a adoré la robe de Lucie. (Principe 1)
  3. Ils ont joué de la musique des années 80 au mariage. (Principe 2)
  4. Anna n'était pas au mariage de Lucie. (Principe 3)
- (32)

### OPI2mariage (34)

This is a conversation between two friends, Rémy (male speaker) and Anna (female speaker). They are talking about a wedding Anna attended.

Rémy: How was Lucie's wedding this past weekend?

Anna: It was fun. We had a great time.

Rémy: And, did you like her dress?

Anna: If we were in the 80s.

(27)

1. Anna did not like Lucie's wedding dress very much (CORRECT)
  2. Anna really liked Lucie's dress at the wedding. (Principe 1)
  3. They played 80s music at Lucie's wedding party. (Principe 2)
  4. Anna did not go to Lucie's wedding. (Principe 3)
- (32)

### OPI3lave-vaisselle—Q35 [used as practice 2]

Conversation entre deux colocataires, Alexandre (homme) et Laura (femme). Ils discutent de ce qu'ils devraient faire avec leur lave-vaisselle qui est cassé.

Alexandre: Hé, Laura. Est-ce que tu as vu que le lave-vaisselle était cassé hier soir?

Laura: Oui, j'ai remarqué. Je m'en suis rendu compte ce matin.

Alexandre: Qu'est-ce qu'on devrait faire à ton avis? Je peux essayer de le réparer moi-même.

Laura: La dernière fois que tu l'as réparé, il était encore cassé le jour suivant.

(52)

1. Laura ne pense pas qu'Alexandre devrait réparer le lave-vaisselle (CORRECT)
  2. Laura pense que c'est une bonne idée qu'Alexandre répare le lave-vaisselle. (Principe 1)
  3. C'est la première fois que le lave-vaisselle de Laura et Alexandre est cassé. (Principe 2)
  4. Alexandre and Laura répareront le lave-vaisselle ensemble. (Principe 3)
- (40)

### OPI3lave-vaisselle (35)

This is a conversation between two roommates, Alexandre (male speaker) and Laura (female speaker). They are talking about what to do with the broken dishwasher.

Alexandre: Hey, Laura. Did you see that the dishwasher broke last night?

Laura: Yes, I noticed. I realized this morning.

Alexandre: What do think we could do? I can try to fix it by myself.

Laura: Last time you fixed it, it broke the next day.

(42)

1. Laura does not think Alexandre should fix the dishwasher. (CORRECT)
2. Laura thinks it is a good idea that Alexandre fixes the dishwasher. (Principle 1)
3. It is the first time that Alexandre and Laura's dishwasher breaks. (Principle 2)
4. Alexandre and Laura are fixing the dishwasher together. (Principle 3)

(40)

### OPI4examen - Q36

Conversation entre deux étudiants, Abdel (homme) et Clara (femme). Ils discutent de leur examen de Chimie

Abdel: Clara, comment ça va?

Clara: Très bien merci. Tu as réussi l'examen de chimie?

Abdel: Je ne l'ai pas encore passé. J'avais un rendez-vous chez le docteur à la même heure donc je n'ai pas pu aller en cours. C'était difficile?

Clara: Je n'ai même pas compris les instructions.

(46)

1. Clara pense que l'examen de chimie était difficile. (CORRECT)
2. Clara n'a eu aucun problème avec l'examen de chimie. (Principle 1)
3. Clara pense que le prof donne de bonnes instructions. (Principle 2)
4. Clara et Abdel ont passé l'examen de chimie ensemble. (Principle 3)

(35)

### OPI4examen (36)

This is a conversation between two classmates, Abdel (male speaker) and Clara (female speaker). They are talking about their Chemistry exam.

Abdel: Clara, how have you been?

Clara: Very good, thanks. How did you do on the chemistry exam?

Abdel: I haven't taken it yet. I had a doctor's appointment at the same time so I couldn't go to class. Was it very difficult?

Clara: I didn't even understand the instructions.

(46)

1. Clara thinks the chemistry exam was very difficult. (CORRECT)
2. Clara did not have a hard time completing the chemistry exam. (Principle 1)
3. Clara thinks the instructor provides good instructions. (Principle 2)
4. Clara and Abdel took the Chemistry test together. (Principle 3)

(34)

### OPI5robe - Q37

Conversation entre deux amis, Thomas (homme) et Coralie (femme). Ils parlent de la couleur de la robe que Coralie a achetée.

Coralie: Thomas, j'ai fait du shopping hier et je me suis achetée une jolie robe bleue.  
Thomas: Ah bon, toi, faire du shopping! C'est étrange!  
Coralie: Attends je vais te montrer. Est-ce que tu trouves que cette couleur me va bien?  
Thomas: Le jaune te va mieux.  
(43)

1. Thomas ne pense pas que le bleu aille très bien à Coralie. (CORRECT)
  2. Thomas pense que la robe bleue sied bien à Coralie. (Principe 1)
  3. Thomas pense que la couleur jaune n'irait pas du tout à Coralie. (Principe 2)
  4. Thomas et Coralie ont fait du shopping ensemble hier. (Principe 3)
- (34)

### OPI5robe (37)

This is a conversation between two friends, Thomas (male speaker) and Coralie (female speaker). They are talking about the color of the dress that Coralie bought.

Coralie: Thomas, I went shopping yesterday and I bought a beautiful blue dress.  
Thomas: You going shopping? That's so weird!  
Coralie; Let me show you. Do you like how the color looks on me?  
Thomas: You'd look better in yellow.  
(36)

1. Thomas does not like how the blue color looks on Coralie. (CORRECT)
  2. Thomas thinks Coralie looks good in the blue dress. (Principe 1)
  3. Thomas thinks the yellow color would look worse on Coralie. (Principe 2)
  4. Thomas and Coralie went shopping together. (Principe 3)
- (36)

### OPI6coiffure - Q38

Conversation entre deux amis, Grégoire (homme) et Melissa (femme). Ils discutent de la nouvelle coupe de cheveux de Melissa

Melissa: Salut, Grégoire! Ça faisait longtemps! Tu as l'air en forme!  
Grégoire: Merci, Melissa! Tu as une nouvelle coupe de cheveux?  
Melissa: Oui, j'ai été chez le coiffeur la semaine dernière! Qu'est-ce que tu en penses?  
Grégoire: Je préférerais la coupe que tu avais avant.  
(41)

1. Grégoire n'aime pas la nouvelle coupe de Melissa. (CORRECT)
  2. Grégoire pense que cette nouvelle coupe va bien à Melissa. (Principe 1)
  3. Grégoire a une nouvelle coupe de cheveux. (Principe 2)
  4. Grégoire et Melissa ont été ensemble chez le coiffeur la semaine dernière. (Principe 3)
- (37)

### OPI6coiffure (38)

This is a conversation between two friends, Grégoire (male speaker) and Melissa (female speaker). They are talking about the new haircut Melisa got.

Melissa: Hi, Grégoire! Long time no see! You look great!

Grégoire: Thank you very much, Melisa! Did you get a new hair cut?

Melissa: Yes! I went to the hair salon last week! What do you think about my new hair style?

Grégoire: I preferred the previous haircut.

(44)

1. Grégoire does not like Melissa's new haircut. (CORRECT)
2. Grégoire thinks that the new haircut looks good on Melissa. (Principle 1)
3. Grégoire just got a new haircut. (Principle 2)
4. Grégoire and Melissa went together to the hairdresser's last week to get their hair cut. (Principle 3)

(38)

### OPI7massage - Q39

Conversation entre deux amis, Gauthier (homme) et Sophia (femme). Ils parlent du massage que Sophia a reçu au spa.

Gauthier: Salut, Sophia! J'espère que tu t'es bien reposée au spa le week-end dernier.

Sophia: J'ai passé tout mon temps dans le sauna.

Gauthier: Ah et ton massage?

Sophia: Je n'avais pas mal au dos avant ce massage!

(34)

1. Sophia pense que le massage n'était pas bon. (CORRECT)
2. Sophia a beaucoup apprécié son massage du dos. (Principle 1)
3. Le dos de Sophia ne lui fait pas mal maintenant. (Principle 2)
4. Sophia and Gauthier ont été au sauna ensemble ce week-end. (Principle 3)

(36)

### OPI7massage (39)

This is a conversation between two friends, Gauthier (male speaker) and Sophia (female speaker). They are talking about the massage Sofia got at the spa.

Gauthier: Hi, Sophia! I hope you could get to relax last weekend at the spa.

Sophia: I spent most of my time at the sauna.

Gauthier: And, what about the back massage you got?

Sophia: My back hurts more before!

(36)

1. Sophia believes the massage for the back was not good. (CORRECT)
2. Sophia liked the massage she got on her back. (Principle 1)
3. Sophia's back does not hurt right now. (Principle 2)
4. Sophia and Gauthier went together to the sauna last weekend. (Principle 3)

(36)

### OPI8film – Q40

Conversation entre deux amis, Marc (homme) et Florence (femme). Ils parlent du film que Florence a vu le week-end dernier

Marc: Florence, tu n'es pas venue au diner chez Marianne samedi dernier. Tu nous as manqué!

Florence: Ouais je pouvais pas y aller. Je suis allée au cinéma avec mes colocataires.

Marc: Comment était le film?

Florence: Je me suis endormie pendant les 15 premières minutes.

(42)

1. Florence pense que le film était très ennuyeux. (CORRECT)
2. Florence a beaucoup aimé le film qu'elle a vu. (Principle 1)
3. Florence veut dormir pendant 15 minutes. (Principle 2)
4. Florence et Marc ont diné ensemble samedi dernier. (Principle 3)

(31)

### OPI8film (40)

This is a conversation between two friends, Marc (male speaker) and Florence (female speaker). They are talking about the movie Florence watched last weekend.

Marc: Florence, you didn't come to the dinner at Marianne's last Saturday. We missed you!

Florence: Yeah, I couldn't come. I went to the movies with my roommates.

Marc: And how was the movie?

Florence: I fell asleep within the first 15 minutes.

(39)

1. Florence thinks the movie was very boring. (CORRECT)
2. Florence really liked the movie she saw with her roommates. (Principle 1)
3. Florence wants to sleep for 15 minutes. (Principle 2)
4. Florence and Marc had dinner together last Saturday. (Principle 3)

(32)

### OPI9entretien – Q41

Conversation entre deux amis, Julien (homme) et Amélie (femme). Ils discutent de l'entretien d'embauche qu'Amélie a passé la semaine dernière.

Julien: Salut, Amélie! Ça va? Quand est-ce que tu as ton entretien?

Amélie: Mon entretien? C'était lundi dernier.

Julien: Oh, je savais pas que c'était la semaine dernière! Et alors? Comment ça s'est passé?

Amélie: Je ne pense pas que je vais avoir le poste.

(41)

1. Amélie pense qu'elle n'a pas réussi son entretien. (CORRECT)
2. Amélie pense qu'elle a bien réussi son entretien. (Principle 1)
3. Amélie a déjà décroché l'emploi pour lequel elle a passé l'entretien. (Principle 2)



4. Amélie et Julien ont passé un entretien le mois dernier. (Principe 3)  
(37)

### OPI9entretien (41)

This is a conversation between two friends, Julien (male speaker) and Amélie (female speaker). They are talking about the job interview Amélie had last week.

Julien: Hi, Amélie! How are you? When is your job interview?

Amélie: Hi, Julien! I'm fine, and you? My job interview was last Monday.

Julien: Oh, I didn't know it was last week! And how was it?

Amélie: I don't think I'm gonna get the job.

(42)

1. Amélie thinks she did not do well on the job interview. (CORRECT)
2. Amélie thinks she did very well on the job interview. (Principe 1)
3. Amélie has already been given the job. (Principe 2)
4. Amélie and Julien had a job interview last month. (Principe 3)

(37)

### OPI10examen - Q42

Conversation entre deux étudiants, Alain (homme) et Lucille (femme). Ils discutent de leur examen de français

Lucille: Oh, Alain. Je suis contente d'avoir fini l'examen de français. J'étais super nerveuse.

Alain: Oui, Lucille. Je suis d'accord. C'était difficile et honnêtement je sais pas si j'ai réussi.

Lucille: Moi je pense que j'avais les bonnes réponses. Pas toi?

Alain: Je n'ai pas répondu à 10 questions sur 15.

(47)

1. Alain est persuadé qu'il n'a pas bien réussi à l'examen. (CORRECT)
2. Alain est convaincu qu'il a très bien réussi à l'examen. (Principe 1)
3. Alain et Lucille pensent qu'ils vont avoir un 10 à l'examen (Principe 2)
4. Alain and Lucille ont beaucoup étudié pour leur examen de français. (Principe 3)

(42)

### OPI10examen (42)

This is a conversation between two classmates, Alain (male speaker) and Lucille (female speaker). They are talking about how they did in their French test.

Lucille: Oh, Alain. I'm glad we already finished our French test. I was so nervous.

Alain: Yes, Lucille. You are right. It was difficult. I don't know how I did.

Lucille: I think I had the right answers. You don't?

Alain: I left 10 out of 15 questions unanswered.

(45)

1. Alain is convinced that he did badly on the test. (CORRECT)
2. Alain is convinced that he got a good grade on the test. (Principe 1)

3. Alain and Lucille think they got a 10 on the test. (Principle 2)
  4. Alain and Lucille studied very hard for their French test. (Principle 3)
- (43)

### OPI11meubles – Q43

Conversation entre deux colocataires, Arnaud (homme) et Camille (femme). Ils discutent de la disposition de certains meubles dans leur appartement

Camille: Arnaud, Je suis désolée de t'embêter mais est-ce que tu pourrais m'aider s'il te plait? Je voudrais réorganiser les meubles dans ma chambre.

Arnaud: Oui pas de problème!

Camille: Ok mettons le bureau ici à gauche en face du mur.

Arnaud: Je pense que tu aurais plus de lumière si on le mettait sous la fenêtre.

(53)

1. Selon Arnaud, mettre le bureau en face du mur n'est pas une bonne idée. (CORRECT)
  2. Arnaud pense que le bureau en face du mur est une bonne idée. (Principle 1)
  3. Camille veut mettre son lit sous la fenêtre. (Principle 2)
  4. Camille et Arnaud sont dans la chambre d'Arnaud. (Principle 3)
- (43)

### OPI11meubles (43)

This is a conversation between two roommates, Arnaud (male speaker) and Camille (female speaker). They are talking about where to place certain pieces of furniture.

Camille: Arnaud, I'm sorry to bother you but could you help me, please? I'm reorganizing my furniture in my room.

Arnaud: Yea no problem!

Camille: Let's move the desk towards the left, facing the wall.

Arnaud: I think you will have more light if we move it under the window.

(46)

1. Arnaud does not think moving the desk facing the wall is a good idea. (CORRECT)
  2. Arnaud likes the idea of moving the desk facing the wall. (Principle 1)
  3. Camille wants to place her bed under the window. (Principle 2)
  4. Camille and Arnaud are in Arnaud's bedroom. (Principle 3)
- (41)

### OPI12musique – Q44

Conversation entre deux colocataires, Jérôme (homme) et Samantha (femme). Ils discutent du groupe de music techno préféré de Jérôme.

Jérôme: Hé, Samantha! Mon groupe préféré de musique techno est en concert en ville Samedi.

Samantha: Ah bon? Tu dois être super content!

Jérôme: Oui carrément! Je suis même plus que super content! C'est le meilleur groupe de techno en ce moment.

Samantha: Ah ça c'est ton opinion!

(44)

1. Samantha pense que le groupe préféré de Jérôme n'est pas bon. (CORRECT)
2. Samantha pense que le groupe préféré de Jérôme est excellent. (Principe 1)
3. Samantha voudrait aller voir le groupe en concert. (Principe 2)
4. Samantha et Jérôme jouent dans un groupe de techno ensemble. (Principe 3)

(39)

### OPI12musique (44)

This is a conversation between two roommates, Jérôme (male speaker) and Samantha (female speaker). They are talking about Jérôme's favorite techno music band.

Jérôme: Hey, Samantha! My favorite techno music band plays in town this Saturday.

Samantha: Really? You must be super excited!

Jérôme: Of course! I'm more than excited! It's the best techno band of the moment.

Samantha: That's what you think.

(36)

1. Samantha thinks that Jérôme's favorite band is not good. (CORRECT)
2. Samantha thinks that the techno music band Jérôme wants to see is excellent. (Principe 1)
3. Samantha thinks she should go see the band play. (Principe 2)
4. Samantha and Jérôme play together in a band. (Principe 3)

(39)

### OPI13bébé—Q45

Conversation entre un couple, Sébastien (homme) et Lola (femme). Ils parlent du cadeau qu'ils vont acheter pour le bébé de leurs amis.

Sébastien: Lola, Antoine m'a dit qu'on pouvait venir voir leur bébé Samedi après-midi.

Lola: Génial! Et si on lui prenait un pyjama? Amélie m'a dit que le bébé avait besoin de pyjamas.

Sébastien: OK!

*(au magasin)*

Qu'est-ce que tu penses de ce pyjama bleu?

Lola: Le vert là lui irait mieux je pense.

(47)

1. Lola n'aime pas vraiment le pyjama bleu. (CORRECT)
2. Lola préfère le pyjama bleu par rapport au pyjama vert. (Principe 1)
3. Une robe verte irait certainement mieux au bébé. (Principe 2)
4. Sébastien et Lola ont été acheter une paire de chaussons pour le bébé. (Principe 3)

(38)

### OPI13bébé (45)

This is a conversation between a couple, Sébastien (male speaker) and Lola (female speaker). They are talking about the gift they will buy for their friends' newborn baby.

Sébastien: Lola, Antoine told us we could visit their baby on Saturday afternoon.

Lola: Fantastic! What about buying him some pajamas? Amélie told me the baby needs pajamas.

Sébastien: Yes!

*(At the store)*

What about this blue one?

Lola: This green one would look better on him.

(40)

1. Lola does not like the pajama in blue. (CORRECT)
2. Lola likes the pajama in blue much better than the green one. (Principle 1)
3. A green dress would fit the baby more nicely. (Principle 2)
4. Sébastien and Lola went to buy a pair of shoes for the baby. (Principle 3)

(42)

### OPI14épicié - Q46

Conversation entre deux colocataires, Christian (homme) et Michaela (femme). Ils parlent des ingrédients dont ils ont besoin pour préparer un poulet pour un dîner avec des amis.

Christian: Qu'est-ce que ta mère utilise quand elle fait du poulet? Son poulet est toujours délicieux!

Michaela: Elle met des rondelles de citrons et des épices avec le poulet dans le four.

Christian: Ah! Je pensais rajouter du Tabasco pour avoir un goût différent.

Michaela: Je suis allergique aux sauces trop épicées!

(48)

1. Michaela ne tolère pas la nourriture épicée. (CORRECT)
2. Michaela veut que Christian rajoute du Tabasco au poulet. (Principle 1)
3. Michaela aime préparer ses repas avec beaucoup d'épices (Principle 2)
4. Michaela, sa mère et Christian préparent du poulet pour un dîner en famille (Principle 3)

(37)

### OPI14épicié (46)

This is a conversation between two roommates, Christian (male speaker) and Michaela (female speaker). They are talking about the ingredients needed to bake some chicken for their dinner with some friends.

Christian: What does your mother use when she makes chicken? Her chicken is always so delicious!

Michaela: She adds some sliced lemons and spices before baking it.

Christian: Ah! I thought about adding some Tabasco to give it a different flavor.

Michaela: I am allergic to spicy sauces.

(44)

1. Michaela cannot eat spicy food. (CORRECT)
2. Michaela wants Christian to add Tabasco to the chicken. (Principle 1)
3. Michaela likes preparing all her food with spicy sauces. (Principle 2)
4. Michaela, her mom and Christian are preparing the chicken for a family dinner. (Principle 3)

(36)

### OPI15tableaux – Q 47

Conversation entre deux étudiants, David (homme) et Léa (femme). Ils discutent des nouveaux tableaux électroniques qui viennent d'être installés dans leur salle de classe

David: Léa, tu as vu les nouveaux tableaux électroniques dans la salle de classe?

Léa: Oui, David. Ils ont enlevé les vieux tableaux hier et les ont remplacés avec des nouveaux tôt ce matin.

David: Je préfère les nouveaux! C'est génial de pouvoir écrire sur le tableau depuis l'ordinateur.

Léa: Moi je préfère écrire avec la craie!

(53)

1. Léa préfère les vieux tableaux par rapport aux nouveaux. (CORRECT)
  2. Léa préfère les tableaux électroniques pour pouvoir écrire depuis l'ordinateur. (Principe 1)
  3. Léa pense qu'on se salit vite avec la craie. (Principe 2)
  4. David et Léa n'ont pas été en cours ce matin. (Principe 3)
- (38)

### OPI15tableaux (47)

This is a conversation between two classmates, David (male speaker) and Léa (female speaker). They are talking about the new boards installed in their classroom.

David: Léa, did you see the new electronic boards in class?

Léa: Yes, David. They took away the old ones yesterday and replaced them with new ones early this morning.

David: I like the new ones much better. Writing on the board from the computer is so much fun.

Léa: I like writing with chalk better.

(52)

5. Léa likes the old boards better than the new ones. (CORRECT)
  6. Léa prefers the electronic boards to write from the computer. (Principe 1)
  7. Léa thinks one can get quite dirty using chalk. (Principe 2)
  8. David and Léa did not go to class this morning. (Principe 3)
- (39)

### OPI16lecture – Q48

Conversation entre deux amis, Julien (homme) et Juliette (femme). Ils discutent de leurs préférences en matière de littérature.

Julien: J'ai enfin pu acheter le dernier livre de la série que j'adore.

Juliette: Tu veux dire ceux avec des zombies qui parlent de l'apocalypse?

Julien: Oui! J'aurai fini en deux jours! Est-ce que tu aimes ce genre de romans?

Juliette: Je préfère les histoires romantiques ou les livres qui parlent de voyage.

(49)

1. Juliette ne veut pas lire le livre de Julien sur les zombies. (CORRECT)
2. Juliette voudrait lire le livre sur les zombies. (Principe 1)
3. Juliette aime lire quand elle voyage. (Principe 2)
4. Juliette était avec Julien pour acheter un livre sur les zombies. (Principe 3)

(37)

### OPI16lecture (48)

This is a conversation between two friends, Julien (male speaker) and Juliette (female speaker). They are talking about their reading preferences.

Julien: I could finally buy the last book from the series I like so much.

Juliette: Do you mean those about zombies and the apocalypse?

Julien: Yes! I'll read it in two days. Do you like this type of novels?

Juliette: I prefer the romantic novels or the books about travelling.

(47)

1. Juliette does not feel like reading Julien's novel about zombies. (CORRECT)
2. Juliette wants to read Julien's novels about zombies. (Principe 1)
3. Juliette likes reading when she is travelling. (Principe 2)
4. Juliette went with Julien to buy the zombies novel. (Principe 3)

(34)

### 3. Fillers/distracters/literal meaning

#### Fillers

#### Fill1jeuxvidéo - Q49 - (IRO4)

Conversation entre deux colocataires, Cédric (homme) et Edith (femme). Ils discutent de jeux vidéo

Cédric: Edith, est-ce que tu as des projets pour ce soir ??

Edith: Non pas vraiment pourquoi?

Cédric: J'ai acheté un nouveau jeux vidéo pour ma Nintendo Switch, tu veux jouer avec moi?

Edith: Non merci. Tu me connais, quand quelqu'un joue je préfère toujours aller dans ma chambre pour lire. Je trouve que les jeux vidéo sont vraiment ennuyeux.

(56)

1. Edith pense que les jeux vidéo sont très ennuyeux. (CORRECT)
2. Edith voudrait bien jouer au nouveau jeu vidéo de Cédric. (Principe 1)
3. Edith joue à des jeux vidéo tous les soirs. (Principe 2)
4. Edith et Cédric vont acheter une nouvelle Nintendo Switch ce soir. (Principe 3)

(39)

#### Fill1jeuxvidéo - (IRO4) (49)

This is a conversation between two roommates, Cédric (male speaker) and Edith (female speaker). They are talking about video games.

Cédric: Edith, do you have any plans for tonight?

Edith: Not really, why?

326

Cédric: I just bought a new video game for my Nintendo Switch? Wanna play with me?  
Edith: No thanks. You know me: whenever someone plays, I go to my room to read. I find video games quite boring.  
(47)

1. Edith thinks that video games are pretty boring. (CORRECT)
2. Edith would like to play the new video game with Cédric. (Principle 1)
3. Edith often plays video games in the evening. (Principle 2)
4. Edith and Cédric are going to buy the new Nintendo Switch console tonight. (Principle 3)

(39)

### Fill2quiche - Q50 - (IRO6)

Conversation entre deux amis, Stéphane (homme) et Natacha (femme), au sujet du diner que Natacha a préparé

Natacha: Stéphane, est-ce que tu veux un café après le dessert?  
Stéphane: Ah? Oui bien sûr.  
Natacha: Est-ce que tu as aimé le diner ? Qu'est-ce que tu penses de ma quiche?  
Stéphane: Ton repas était délicieux et j'ai adoré la quiche. J'ai tout mangé!  
(41)

1. Stéphane a beaucoup aimé le diner que Natacha a préparé. (CORRECT)
  2. Stéphane n'a pas apprécié le repas de Natacha pour diner. (Principle 1)
  3. Stéphane pense que la quiche de Natacha n'a pas bon gout. (Principle 2)
  4. Stéphane et Natacha commandent toujours un café au restaurant. (Principle 3)
- (40)

### Fill2quiche - (IRO6) 50

This is a conversation between two friends, Stéphane (male speaker) and Natacha (female speaker). They are talking about the dinner Natacha prepared.

Natacha: Stéphane, would you like some coffee after dessert?  
Stéphane: Why? Yes, of course.  
Natacha: Did you like dinner? What did you think of my quiche?  
Stéphane: Your food was delicious and I loved the quiche. I ate it all!  
(36)

1. Stéphane really liked the food that Natacha prepared for dinner. (CORRECT)
  2. Stéphane did not like the meal that Natacha made for dinner. (Principle 1)
  3. Stéphane thinks that Natacha's quiche tastes absolutely terrible. (Principle 2)
  4. Stéphane and Natacha always order coffee at the restaurant. (Principle 3)
- (38)

### Fill3gâteau - Q51 - (IRO6)

Conversation entre une mère (femme) et son fils, Nicolas (homme) au sujet du gâteau préparé par la mère

Mère: Je suis en train de tester une nouvelle recette de gâteau au chocolat, tu veux goûter?

Nicolas: Ah? Oui bien sûr.

Mère: Alors qu'est-ce tu en dis, tu penses que ça va plaire à ton père?

Nicolas: Oh la la, c'est trop sucré, j'aime pas trop. Je pense que tu devrais mettre moins de sucre!

(52)

1. Nicolas n'aime pas le gâteau que sa mère est en train de préparer. (CORRECT)
  2. Nicolas aime beaucoup le gâteau de sa mère. (Principe 1)
  3. Nicolas pense que son père va beaucoup aimer ce gâteau. (Principe 2)
  4. Nicolas et sa mère vont faire la cuisine ensemble. (Principe 3)
- (40)

### Fill3gâteau - (IRO6) 51

This is a conversation between a mother (female speaker) and her son, Nicolas (male speaker) about the cake the mother baked.

Mother: I'm trying out a new chocolate recipe: do you want to taste it?

Nicolas: Ah yes, sure!

Mother: So what do you think? Think your dad is going to like it?

Nicolas: Uhm it's too sweet, I don't like it! I think you put too much too sugar in it!

(47)

1. Nicolas does not like the cake that his mother is baking. (CORRECT)
  2. Nicolas really likes his mother's cake. (Principe 1)
  3. Nicolas thinks his dad is going to like this cake a lot. (Principe 2)
  4. Nicolas and his mother are going to bake together. (Principe 3)
- (38)

### Fill 4temps - Q52 - (IRO7 - possibly 2)

Conversation entre deux amis, Eric (homme) et Martine (femme). Ils discutent du temps qu'il fait

Martine: Eric tu es prêt? On y va? Je ne veux pas être en retard à mon cours!

Eric: Ah oui, on y va à pied?

Martine: Oh ben je sais pas, tu as vu le temps qu'il fait? Il pleut et il fait froid dehors!

Eric: Ah oui tu as raison! Heureusement que tu me le dis!

(54)

1. Eric n'avait pas remarqué qu'il faisait mauvais aujourd'hui. (CORRECT)
  2. Eric savait qu'il pleuvait avant même que Martine ne le lui dise. (Principe 1)
  3. Eric est toujours heureux d'aller à l'université en bus. (Principe 2)
  4. Martine et Eric vont toujours à l'université à pied ensemble. (Principe 3)
- (39)

### Fill 4temps - (IRO7 - possibly IRO2) 52

This is a conversation between two friends, Eric (male speaker) and Martine (female speaker). They are discussing the weather.



Martine: Eric are you ready? Let's go! I don't want to be late for class!

Eric: Ok ok are we walking there?

Martine: Oh I don't know, did you see the weather? It's raining and it seems cold outside!

Eric: Oh yes, you're right! I'm glad you told me!

(45)

1. Eric had not realized how bad the weather was today. (CORRECT)
  2. Eric knew that it was raining before Martine even told him. (Principle 1)
  3. Eric is always happy to take the bus to the university. (Principle 2)
  4. Martine and Eric always walk together to the university. (Principle 3)
- (41)

### Fill5robe - Q53 - (IRO13+OPI12)

Conversation entre deux amis, Charlie (homme) et Eva (femme). Ils parlent d'une fête à laquelle ils sont allés.

Charlie: Salut, Eva. Tu t'es bien amusé à la fête de Nathalie? Moi j'ai rencontré une fille!

Eva: Ah bon, qui? A quoi elle ressemble?

Charlie: Elle s'appelle Stéphanie. Elle portait une jolie robe bleue à la fête! Tu vois qui?

Eva: Ah oui, tout à fait, d'ailleurs j'adore sa robe!

(47)

1. Eva pense que la robe que Stéphanie portait est belle. (CORRECT)
  2. Eva pense que la robe de Stéphanie est moche. (Principle 1)
  3. Eva s'est achetée une robe pour aller à une fête. (Principle 2)
  4. Charlie et Eva vont rendre visite à Stéphanie. (Principle 3)
- (37)

### Fill5robe - (IRO13+OPI12) 53

This is a conversation between two friends, Charlie (male speaker) and Eva (female speaker). They're talking about a party that they went to.

Charlie: Hi, Eva. Did you have fun at Nathalie's party? I met a girl there!

Eva: Oh really? Who? What does she look like?

Charlie: Her name is Stéphanie. She was wearing a pretty blue dress! Do you know who I'm talking about?

Eva: Oh yes, totally, and I love her dress!

(48)

1. Eva thinks the dress that Stéphanie was wearing is beautiful. (CORRECT)
  2. Eva thinks the dress that Stéphanie was wearing is ugly. (Principle 1)
  3. Eva bought herself a dress to go to a party. (Principle 2)
  4. Charlie and Eva are going to visit Stéphanie. (Principle 3)
- (38)

### Fill6ponctualité - Q54 - (IRO11)

Conversation entre deux amis, Sylvain (homme) et Isabelle (femme). Isabelle arrive au restaurant avec 20 minutes de retard alors que Sylvain l'attendait.

Isabelle : Salut Sylvain, je m'excuse pour ce retard.

Sylvain: Salut, Isabelle. Ça va? Qu'est-ce qui s'est passé?

Isabelle: Oui je suis vraiment désolée. J'avais perdu mes clefs et ensuite j'ai eu des problèmes avec ma voiture.

Sylvain: Oui, j'étais un peu inquiet: d'habitude tu es toujours à l'heure.

(45)

1. Isabelle est normalement toujours à l'heure à ses rendez-vous. (CORRECT)
  2. Sylvain pense qu'Isabelle arrive souvent en retard à ses rendez-vous. (Principe 1)
  3. Sylvain est toujours inquiet quand Isabelle doit prendre le bus. (Principe 2)
  4. Isabelle et Sylvain ont perdu leurs clefs. (Principe 3)
- (36)

### Fill6ponctualité - (IR011) 54

This is a conversation between two friends, Sylvain (male speakers) and Isabelle (female speaker). Isabelle arrives 20 minutes late at a restaurant where Sylvain is waiting for her.

Isabelle: Hello Sylvain, I apologize for being late.

Sylvain: Hi, Isabelle. Everything ok? What happened?

Isabelle: Yes, I'm sorry: First I lost my keys and then I had some problems with my car.

Sylvain: Oh ok, I was a little bit worried: usually you're always on time.

(43)

1. Isabelle is usually on time to her appointments. (CORRECT)
  2. Sylvain thinks Isabelle is often late to her appointments. (Principe 1)
  3. Sylvain always worries when Isabelle has to take the bus. (Principe 2)
  4. Isabelle and Sylvain drove their car to the restaurant. (Principe 3)
- (36)

### Fill7coiffure - Q55 - (OPI6)

Conversation entre deux amis, François (homme) et Anne-Lise (femme). Ils discutent de la nouvelle couleur de cheveux d'Anne-Lise.

Anne-Lise: Salut, François! Comment ça va? Ça faisait longtemps.

François: Oui c'est vrai! Ça va bien. Mais dis-moi tes cheveux étaient blonds avant non?

Anne-Lise: Oui, j'ai décidé de changer de couleur! Qu'est-ce que tu penses du châtain?

François: Ben j'aime pas trop le châtain en fait.

(43)

1. François pense que le châtain ne va pas à Anne-Lise. (CORRECT)
  2. François aime beaucoup la nouvelle couleur de cheveux d'Anne-Lise. (Principe 1)
  3. François a eu les cheveux châtain à un moment donné. (Principe 2)
  4. François et Anne-Lise sont allés chez le coiffeur ensemble hier. (Principe 3)
- (39)

### Fill7coiffure (OPI6) 55

This is a conversation between two friends François (male speaker) and Anne-Lise (female speaker). They're discussing Anne-Lise's new hair color.

Anne-Lise: Hello, François! How is it going? It's been a while.

François: Yes, you're right! I'm doing well. But, wasn't your hair blond before?

Anne-Lise: Yes, I decided to switch color! What do you think of the brown hair?

François: Well I don't really like it actually.

(43)

1. François thinks brown hair does not suit Anne-Lise very well. (CORRECT)
  2. François really likes Anne-Lise's new hair color. (Principle 1)
  3. François has had both brown and blond hair at some point. (Principle 2)
  4. François and Anne-Lise went to the hair salon together yesterday. (Principle 3)
- (38)

### Fill8film - Q56 - (OPI8)

Conversation entre deux amis, Evan (homme) et Emilie (femme). Ils se rencontrent au café et parlent du film qu'ils ont vu au cinéma le week-end précédent.

Evan: Salut Emilie, comment ça va depuis le week-end dernier?

Emilie: Ça va bien et toi? Tu veux un café?

Evan: Oui, merci. Dis-moi tu as aimé le film d'horreur qu'on a vu samedi dernier?

Emilie: Non, franchement je l'ai pas aimé du tout!

(40)

1. Emilie n'a pas aimé le film qu'ils ont vu. (CORRECT)
  2. Emilie a beaucoup aimé le film qu'elle a vu avec Evan. (Principle 1)
  3. Emilie n'aime pas le café du tout. (Principle 2)
  4. Emilie et Marc sont au cinéma. (Principle 3)
- (33)

### Fill8film (OPI8) 56

This is a conversation between two friends, Evan (male speaker) and Emilie (female speaker). They met at a café and are talking about the movie they saw the weekend before.

Evan: Hi Emilie, how have you been since last weekend?

Emilie: I'm good what about you? Do you want a cup of coffee?

Evan: Yes, thank you. By the way, did you like the horror movie we saw on Saturday?

Emilie: No, I did not like it at all!

(46)

1. Emilie did not like the horror movie they saw. (CORRECT)
  2. Emilie really liked the horror movie they saw. (Principle 1)
  3. Emilie does not like coffee at all. (Principle 2)
  4. Emilie and Marc are at the movie theater. (Principle 3)
- (32)

### Fill9musique - Q57 - (OPI12)

Conversation entre deux colocataires, Jérémie (homme) et Carole (femme). Ils discutent de la chanteuse préférée de Carole.

Carole: Hé, Jérémie! Ça va? Je viens enfin de recevoir le dernier album de Céline Dion.

Jérémie: Ah? Tu dois être très contente!

Carole: Oh la la oui! J'étais impatiente! C'est vraiment la meilleure chanteuse du monde, non?

Jérémie: Non, je suis pas du tout d'accord. J'aime vraiment pas sa voix!

(47)

1. Jérémie pense que la chanteuse préférée de Carole n'est pas terrible. (CORRECT)
2. Jérémie pense que la chanteuse préférée de Carole est excellente. (Principe 1)
3. Jérémie n'est jamais d'accord avec Carole sur la musique. (Principe 2)
4. Jérémie et Carole vont aller voir Céline Dion en concert. (Principe 3)

(40)

### Fill9musique (OPI12) 57

This is a conversation between two roommates, Jérémie (male speaker) and Carole (female speaker). They are talking about Carole's favorite singer.

Carole: Hey, Jérémie! How are you? I just received Céline Dion's last album.

Jérémie: Ah? You must be so excited!

Carole: Yes definitely! I could hardly wait! She's the best singer in the world, isn't she?

Jérémie: No, I completely disagree. I really do not like her voice!

(44)

1. Jérémie thinks that Carole's favorite singer is not that great. (CORRECT)
2. Jérémie thinks that Carole's favorite singer is excellent. (Principe 1)
3. Jérémie always disagrees with Carole when it comes to music. (Principe 2)
4. Jérémie and Carole are going to see Céline Dion in concert. (Principe 3)

(39)

### Fill10café - Q58 - (IRO16)

Conversation entre deux amis, Kévin (homme) et Sylvie (femme). Ils sont en train de commander des boissons dans un café.

Sylvie: Salut, Kévin! Ça faisait longtemps.

Kévin: Oui n'est-ce pas? Je viens de me commander une tasse de thé. Qu'est-ce que tu vas prendre?

Sylvie: Je voudrais une tasse de café avec du chocolat et de la sauce au caramel.

Kévin: Et ben dis donc, je vois que tu aimes bien le café bien sucré.

(51)

1. Kévin pense que Sylvie a commandé une tasse de café très sucré. (CORRECT)
2. Kévin pense que Sylvie n'aime pas le café sucré. (Principe 1)
3. Kévin a commandé une tasse de café. (Principe 2)
4. Kévin et Sylvie aiment tous les deux le thé. (Principe 3)

(37)

### Fill10café (IRO16) 58

This is a conversation between two friends, Kévin (male speaker) and Sylvie (female speaker). They are ordering drinks at a coffee shop

Sylvie: Hi, Kévin! It's been a while.

Kévin: I know, right? I just ordered a cup of tea for myself. What would you like to drink?

Sylvie: I'll have a cup of coffee with chocolate powder and caramel sauce.

Kévin: Wow, I can see that you like your coffee really sweet.

(47)

5. Kévin thinks Sylvie ordered a very sweet cup of coffee. (CORRECT)
6. Kévin thinks Sylvie does not like her coffee with too much sugar. (Principle 1)
7. Kévin ordered a cup of black coffee. (Principle 2)
8. Kévin and Sylvie both like tea. (Principle 3)

(35)

## Appendix D. Dialogue, video and MCQ lengths

**Table 26**

*Dialogue, video and MCA lengths*

Q#	Item	Dialogue in French - word count	Video time (in seconds)	MCQ in English - word count	MCQ in French - word count
<b>Indirect Refusals (IRF)</b>					
1	IR1repas	41	13	37	37
2	IRF2caf��teria	50	19	38	38
3	IRF3voyage	47	14	39	39
4	IRF4r��visions	43	13	41	41
5	IRF5d��jeuner	41	13	37	37
6	IRF6devoir	42	15	40	40
7	IRF7supermarch��	47	13	35	35
8	IRF8exercice	47	13	35	35
9	IRF9buffet	35	11	37	37
10	IRF10pluie	45	10	38	38
11	IRF11dessert	50	14	36	36
12	IRF12notes	45	12	36	36
13	IRF13chaussures	54	12	41	41
14	IRF14snacks	41	12	39	39
15	IRF15vacances	40	12	41	41
16	IRF16mus��e	40	12	40	40
<b>Irony (IRO)</b>					
17	IRO1volume	41	24	38	38
18	IRO2climat	39	12	38	38
19	IRO3devoirs	48	17	37	37
20	IRO4t��l��vision	55	16	39	39
21	IRO5Europe	52	17	36	36
22	IRO6diner	43	14	38	38
23	IRO7chaleur	47	15	41	41
24	IRO8grandes��coles	51	14	39	39
25	IRO9corrections	49	13	41	41
26	IRO10entorse	47	18	34	35
27	IRO11ponctualit��	42	19	36	36
28	IRO12matin	52	18	41	41
29	IRO13f��te	56	18	38	38
30	IRO14froid	46	13	39	39
31	IRO15meubles	51	16	38	41
32	IRO16capuccino	61	18	35	34



		mean word count IRF	average time IRF	mean MCQ word count IRF	mean MCQ word count IRF
		<b>44.25</b>	<b>13.00</b>	<b>38.13</b>	<b>38.13</b>
	<i>SD</i>	4.80	2.00	2.09	2.09
	range	35-54	10-19	35-41	35-41
		mean word count IRO	average time IRO	mean MCQ word count IRO	mean MCQ word count IRO
		<b>48.75</b>	<b>16.38</b>	<b>38.00</b>	<b>38.19</b>
	<i>SD</i>	5.89	2.96	2.07	2.20
	range	39-56	12-24	34-41	34-41
		mean word count OPI	average time OPI	mean MCQ word count OPI	mean MCQ word count OPI
		<b>44.56</b>	<b>13.44</b>	<b>37.44</b>	<b>37.38</b>
	<i>SD</i>	6.35	2.66	3.39	3.26
	range	30-53	9-19	32-43	31-43
		mean word count fillers	average time fillers	mean MCQ word count fillers	mean MCQ word count fillers
		<b>47.60</b>	<b>14.00</b>	<b>37.40</b>	<b>37.80</b>
	<i>SD</i>	5.50	1.83	2.50	2.25
	range	40-56	11-17	32-43	33-40



## Appendix E. Opinion/judgement survey

### Instructions:

1) Sur une échelle de 1 à 5, indiquez si les scénarios/dialogues du document ci-joint sont crédibles/plausibles en encerclant le chiffre correspondant. Si possible prenez des notes pour discuter de changements possibles.

2) Indiquez si la dernière phrase du dialogue correspond à l'affirmation ci-dessous (oui ou non selon si vous êtes d'accord).

3) Indiquez si vous pensez que la dernière phase du dialogue est un exemple de refus indirect/ironie/opinion indirect/opinion directe (selon le dialogue, la catégorie sera différente)

### Dialogue 1 à 16 : exemples de personnes qui refusent de façon indirecte

#### Dialogue 1

1) Selon vous, ce dialogue/scénario est...

pas du tout crédible/plausible	peu crédible/plausible	pas sûr	assez crédible/plausible	tout à fait crédible/plausible
1	2	3	4	5

2) → Cécile ne veut pas dîner avec François ce soir. Oui Non

3) Je pense que ce dialogue est un exemple de personne qui refuse de façon indirecte

pas du tout	pas vraiment	pas sûr	plutôt	tout à fait
1	2	3	4	5

#### Dialogue 2

1) Selon vous, ce dialogue/scénario est...

pas du tout crédible/plausible	peu crédible/plausible	pas sûr	assez crédible/plausible	tout à fait crédible/plausible
1	2	3	4	5

2) → Marie ne veut rien à manger. Oui Non

3) Je pense que ce dialogue est un exemple de personne qui refuse de façon indirecte

pas du tout	pas vraiment	pas sûr	plutôt	tout à fait
1	2	3	4	5

### Dialogue 3

1) Selon vous, ce dialogue/scénario est...

pas du tout crédible/plausible	peu crédible/plausible	pas sûr	assez crédible/plausible	tout à fait crédible/plausible
1	2	3	4	5

2) → Marie n'a pas assez d'argent pour aller visiter Londres. Oui Non

3) Je pense que ce dialogue est un exemple de personne qui refuse de façon indirecte

pas du tout	pas vraiment	pas sûr	plutôt	tout à fait
1	2	3	4	5

## Appendix F. Reviewers dialogue ratings

**Table 27**

*Reviewers dialogue ratings*

Dialogue #	Plausibility/credibility of the dialogues/situations				Does the dialogue match the indirect meaning category?			
	reviewer 1	reviewer 2	reviewer 3	reviewer 4	reviewer 1	reviewer 2	reviewer 3	reviewer 4
1	5	4	5	5	5	5	5	5
2	5	5	5	5	5	5	5	5
3	4	3	4	4	4	4	4	4
4	5	5	5	5	5	5	5	5
5	5	4	4	5	5	5	5	5
6	5	5	5	5	5	5	5	5
7	5	5	5	5	5	5	5	5
8	5	5	5	5	5	5	5	5
9	5	5	5	5	5	5	5	5
10	5	4	5	4	5	4	4	5
11	5	4	4	5	5	4	4	5
12	5	5	5	5	5	5	5	5
13	4	3	4	5	4	4	4	4
14	5	4	5	5	5	5	5	5
15	5	5	5	5	5	5	5	5
16	5	5	5	5	5	4	4	5
17	4	4	4	5	5	4	5	4
18	4	3	4	4	4	4	5	4
19	5	5	4	5	5	5	4	5
20	4	4	4	5	4	4	4	5
21	5	5	5	5	5	5	5	5
22	5	4	4	4	4	4	4	4
23	5	5	5	5	5	5	5	5
24	5	5	3	5	5	5	3	5
25	5	5	5	5	5	5	5	5
26	5	5	5	5	5	5	5	5
27	5	5	5	5	5	5	5	5
28	4	3	3	4	4	3	3	4
29	4	4	4	4	5	4	4	5
30	5	5	5	5	5	5	5	5
31	4	3	4	4	5	4	4	5
32	5	4	5	5	5	5	5	5
33	5	5	5	5	5	4	5	5

34	5	5	5	5	5	5	4	4
35	5	5	5	5	5	5	4	5
36	5	5	5	5	5	4	5	4
37	5	5	5	5	5	5	5	5
38	5	5	5	5	5	5	5	5
39	5	5	4	5	5	5	4	5
40	5	5	4	5	5	5	4	5
41	5	5	5	5	4	4	4	4
42	5	5	4	5	5	4	4	5
43	5	5	5	5	4	3	4	4
44	5	5	5	5	5	4	5	4
45	5	5	5	5	4	4	5	4
46	5	5	5	5	5	4	5	5
47	5	5	5	5	5	5	5	5
48	5	5	5	5	5	5	5	5
49	5	5	5	5	5	5	5	5
50	5	5	5	5	5	5	5	5
51	5	5	5	5	5	5	5	5
52	5	5	5	5	5	4	5	5
53	5	5	5	5	5	5	5	5
54	5	5	5	5	5	5	5	5
55	5	4	5	5	5	4	5	5
56	5	5	5	5	5	5	5	5
57	5	4	5	5	5	5	5	5
58	5	5	5	5	5	4	4	5

mean	4.76
SD	0.49

mean	4.71
SD	0.49

## Appendix G. Biographical data collection

Full Name: \_\_\_\_\_

Age (in years): \_\_\_\_\_

Native/First language:

- English
- Other (specify): \_\_\_\_\_

➔ If other: How long have you lived in the U.S.? (in years): \_\_\_\_\_

How many years have you studied/been speaking English? \_\_\_\_\_

How long have you been studying French (in semesters or years—indicate whether it was in High School, college, etc.)

Have you studied abroad or lived in a French-Speaking country?

- No
- Yes: how long and where? \_\_\_\_\_

Gender identification:

- Male
- Female
- Other

Best email to reach you, should you be selected for a follow-up interview: \_\_\_\_\_

## Appendix H. Test instructions (L2 – English version)

### Instructions:

**You will be asked to watch a series of video clips. Please only watch each video ONCE.**

**After each video, you will get a multiple-choice question with four possible answers. Select the appropriate statement based on your understanding of the last statement in the video.**

**The computer will also record the time you take to answer that question from the moment the questions appear on the screen and the time you click on your answer. It is therefore important that you do not get distracted and click on the answer you think is correct as soon as you have it. However, we also want you to take the time to read all the possible answers and if you need to think about it, it's ok. This is not a speed competition: so take the time that you need to answer as long as you stay focused on the task.**

## Appendix I. Pilot Study recruitment message

Message that was sent out to French students at two French public universities to recruit native speakers for our pilot study.

*Bonjour,*

*Actuellement étudiante en doctorat à l'université de Montpellier, je m'intéresse à la compréhension de certains aspects de la conversation en français par les apprenants américains.*

*Je recherche **des étudiants de langue française** pour la **version pilote** de mon projet.*

*Le questionnaire en ligne comprend une série de courtes vidéos en français, suivies de questions à choix multiples simples et faciles pour des locuteurs natifs.*

*La durée anticipée de l'enquête est d'environ 35 minutes, mais chaque vidéo dure seulement une douzaine de secondes environ.*

*Pour accéder au questionnaire en ligne, cliquez sur ce lien: [https://unt.az1.qualtrics.com/jfe/form/SV\\_6gtDAGSYVh02GS9](https://unt.az1.qualtrics.com/jfe/form/SV_6gtDAGSYVh02GS9)*

*Je vous remercie d'avance pour votre participation, sans laquelle je ne pourrais compléter cette étape primordiale de ma thèse. Si vous avez des questions sur ce projet, vous pouvez me contacter à l'adresse email suivante: [laetitia.knight@gmail.com](mailto:laetitia.knight@gmail.com)*

*Cordialement,*

*Laetitia Knight*

## Appendix J. ACTFL Can-Do Statements

### ACTFL CAN-DO STATEMENTS

*On a scale of 1 to 5, with 1 (on the left) being "I can't do this/it's too difficult" and 5 meaning "I can do this/No problem", evaluate the following statements based on what you think are you able to do in FRENCH.*

I can read movie reviews to choose what to watch.

I can understand the main points of a science fiction short story.

I can understand short presentations about famous people in history.

I can understand the main points of recorded short story

I can interact with the hotel staff to request a room switch.

I can exchange messages with a friend to explain why I had to cancel a planned activity.

I can exchange advice online about how to behave when visiting a family for dinner in a different culture.

I can add my opinion about a social issue onto an online blog or social media.

I can make a presentation describing the highlights of a recent or upcoming family event.

I can write about a social, cultural or political event that occurred or will occur in my community.

I can promote an art or music event by presenting and elaborating on its social or artistic value.

I can understand some events described in an excerpt from an historical journal.

I can explain the process for updating a website or social media account.

I can draft and revise a synopsis or abstract for a project, research study, or conference.

I can follow the chronology of a novel's plot.



## Appendix K. Informed Consent Form



UNIVERSITY OF NORTH TEXAS®

### Informed Consent for Studies with Adults

**TITLE OF RESEARCH STUDY:** Indirect Speech Comprehension in French

**RESEARCH TEAM:** Laetitia Knight, University of North Texas (UNT), Department of World Languages, Literatures, and Cultures, 940-565-4404, [laetitia.knight@unt.edu](mailto:laetitia.knight@unt.edu)

You are being asked to participate in a research study. Taking part in this study is voluntary. The investigator will explain the study to you and will answer any questions you might have. It is your choice whether or not you take part in this study. If you agree to participate and then choose to withdraw from the study, that is your right, and your decision will not be held against you.

You are being asked to participate in a research study, which involves determining the extent of which learners of French are able to detect and understand various occurrences of non-literal language. The purpose of the study is to identify the difficulties learners of French may have in comprehending different types of indirect meaning.

Your participation in this research study involves an assignment that all students will have to complete as part of their normal coursework. During this activity, you will watch short video clips and answer a questionnaire. You may not opt out of this activity--and will receive a completion grade for it--but you have to consent for the researcher to use the anonymized data for her research. Based on your answers, you may also be selected for a short (less than 30 minutes) follow-up interview with the researcher. You will have the option to decline the interview. More details will be provided in the next section.

There are no foreseeable risks in taking part in this study.

**DETAILED INFORMATION ABOUT THIS RESEARCH STUDY:** The following is more detailed information about this study, in addition to the information listed above.

**PURPOSE OF THE STUDY:** The purpose of this replication study is to determine the extent of which learners of French are able to detect and understand various occurrences of non-literal language, and corroborate previous findings. To such aim, the investigator will attempt to identify the difficulties learners of French may have in comprehending different types of indirect meaning. The overarching research question is the following: *Do L2 French learners demonstrate different comprehension accuracy and response speeds across different types of indirect meaning?*

We are hoping that the analysis of the collected data will inform best practices for teaching indirect meaning in French.

**TIME COMMITMENT:** Participation in the activity is expected to last approximately one hour and will be conducted as part of a regularly scheduled in-class assignment. Students who are selected and agree to a follow-up interview will meet with the research for approximately 20-30 minutes.

**STUDY PROCEDURES:**

The activity you will complete is part of your regular classwork. All students will receive a completion grade for completing the assignment. The activity consists of a computer-based survey, which presents a series of video clips in French, followed by multiple-choice comprehension questions in English, that you will need to answer.

Approximately three weeks after this activity, you may be invited to participate in a follow-up audio-recorded interview. Should you not wish to be contacted for the follow-up interview, you may check the appropriate box below. During the interview, the researcher will go over some of the answers and ask participants to explain their thought-process. The researcher's intent will be to identify specific strategies used by the participants. Interviews will last less than 30 minutes each.

**POSSIBLE BENEFITS:** This study is not expected to be of any direct benefit to the participants, but we hope to learn more about the ways learners of French understand indirect meaning in order to inform best practices for teaching French, which may benefit future learners of French.

Participants who are selected and agree to a follow-up interview will receive a tutoring session (should they want it).

**POSSIBLE RISKS/DISCOMFORTS:** This research study is not expected to pose any additional risks beyond what you would normally experience in your regular everyday life. However, if you do experience any discomfort, please inform the research team

**COMPENSATION:** There is no financial compensation.

Participants will be required to complete the activity as part of their regularly scheduled coursework and as such will receive a completion grade. The decision to opt out of the study will only impact whether participants' collected data will be used for this study or not.

As an incentive, students who agree to and are selected to participate in the follow-up interview will receive a complimentary tutoring session (should they need it).

**CONFIDENTIALITY:** Efforts will be made by the research team to keep your personal information private, including research study records, and disclosure will be limited to people who have a need to review this information. All paper and electronic data collected from this study will be stored in a secure location on the UNT campus and/or a secure UNT server (OneDrive) for at least three (3) years past the end of this research. Research records will be labeled with a code and the master key linking names with codes will be maintained in a separate and secure location.

The results of this study may be published and/or presented without naming you as a participant. The data collected about you for this study may be used for future research studies that are not described in this consent form. If that occurs, an IRB would first evaluate the use of any information that is identifiable to you, and confidentiality protection would be maintained

While absolute confidentiality cannot be guaranteed, the research team will make every effort to protect the confidentiality of your records, as described here and to the extent permitted by law.

**CONTACT INFORMATION FOR QUESTIONS ABOUT THE STUDY:** If you have any questions about the study you may contact Laetitia Knight at [Laetitia.knight@unt.edu](mailto:Laetitia.knight@unt.edu). Any questions you have regarding your rights as a research subject, or complaints about the research may be directed to the Office of Research Integrity and Compliance at 940-565-4643, or by email at [untirb@unt.edu](mailto:untirb@unt.edu).

**I have read the consent information and agree for my answers to be used for this research study:**

**Yes, I agree**

**No, I do not agree**

**AUDIO/VIDEO/PHOTOGRAPHY:** The recordings will be kept with other electronic data in a secure UNT OneDrive account for the duration of the study. Transcriptions of the audio-recordings may be used in future publications or presentations.

**I agree** to be contacted for a follow-up audio-recorded interview.

**I do not agree** to be contacted for a follow-up audio-recorded interview.

## Appendix L. Learner Accuracy Random Effects

Table 28 presents the random effects for our Learner Accuracy Model (random intercept), and Table 29 the random slope for that same model.

**Table 28**

*Learner Accuracy Model - Random Intercept*

Participant	Random Intercept	Participants	Random Intercept
1	-16.5358930	18	0.7164263
2	-13.6605065	19	4.2307876
3	-9.1876829	20	2.3138632
4	-12.3825569	21	2.3138632
5	-9.1876829	22	3.9113002
6	-9.1876829	23	2.6333506
7	-10.4656325	24	6.1477120
8	-7.2707586	25	4.2307876
9	-4.0758846	26	7.4256616
10	-4.0758846	27	8.7036111
11	-5.6733216	28	9.0230985
12	-5.3538342	29	12.2179725
13	-2.4784477	30	10.9400229
14	-2.7979351	31	12.2179725
15	-3.7563972	32	12.5374599
16	2.3138632	33	13.8154094
17	0.3969389		

**Table 29**

*Learner Accuracy Model - Random Slope*

Participant/ Condition	Random slope	Participant/ Condition	Random slope
1/FILLER	19.59332532	18/FILLER	-3.41980620
1/IRF	-20.36862281	18/IRF	0.24799266
1/IRO	0.74254509	18/IRO	7.28504863

1/OPI	-16.95671689	18/OPI	-3.37715738
2/FILLER	8.72074743	19/FILLER	2.05697007
2/IRF	-1.68556564	19/IRF	4.31735774
2/IRO	-1.68556564	19/IRO	-2.71969823
2/OPI	-19.38482761	19/OPI	0.69220770
3/FILLER	-3.50078086	20/FILLER	-4.76875661
3/IRF	1.57442920	20/IRF	12.97315419
3/IRO	1.57442920	20/IRO	-1.10095774
3/OPI	-9.08777681	20/OPI	-4.72610778
4/FILLER	-0.80288005	21/FILLER	-4.76875661
4/IRF	-9.80178193	21/IRF	-8.13801371
4/IRO	-2.76472596	21/IRO	-1.10095774
4/OPI	0.64717997	21/OPI	16.38506012
5/FILLER	-3.50078086	22/FILLER	-6.11770701
5/IRF	-5.46262677	22/IRF	11.62420379
5/IRO	8.61148517	22/IRO	-2.44990814
5/OPI	-9.08777681	22/OPI	0.96199778
6/FILLER	-3.50078086	23/FILLER	3.40592047
6/IRF	-12.49968273	23/IRF	5.66630815
6/IRO	8.61148517	23/IRO	-8.40780379
6/OPI	-2.05072084	23/OPI	2.04115810
7/FILLER	6.02284662	24/FILLER	8.88269674
7/IRF	-4.38346644	24/IRF	2.69861726
7/IRO	-4.38346644	24/IRO	-4.33843871
7/OPI	-8.00861649	24/OPI	-0.92653279
8/FILLER	3.32494582	25/FILLER	2.05697007
8/IRF	-14.11842322	25/IRF	-2.71969823
8/IRO	-0.04431128	25/IRO	-2.71969823
8/OPI	3.36759464	25/OPI	7.72926367
9/FILLER	0.62704501	26/FILLER	-0.64093074
9/IRF	-9.77926806	26/IRF	1.61945693
9/IRO	-2.74221209	26/IRO	-5.41759903
9/OPI	7.70674980	26/OPI	12.06841883
10/FILLER	0.62704501	27/FILLER	-10.16455823
10/IRF	4.29484388	27/IRF	7.57735258
10/IRO	-9.77926806	27/IRO	7.57735258
10/OPI	0.66969383	27/OPI	3.95220253
11/FILLER	1.97599541	28/FILLER	-1.98988114
11/IRF	5.64379428	28/IRF	14.34461847
11/IRO	-15.46737362	28/IRO	-6.76654944
11/OPI	2.01864424	28/OPI	3.68241245
12/FILLER	10.15067249	29/FILLER	-4.68778195
12/IRF	-22.77421967	29/IRF	-2.42739428
12/IRO	12.41106017	29/IRO	11.64671766
12/OPI	-5.28820181	29/OPI	8.02156761
13/FILLER	-0.72190540	30/FILLER	4.83584553
13/IRF	2.94589347	30/IRF	12.72587798
13/IRO	-4.09116250	30/IRO	-15.42234589
13/OPI	-0.67925657	30/OPI	9.10072794
14/FILLER	-8.89658248	31/FILLER	-4.68778195
14/IRF	3.21568355	31/IRF	11.64671766

14/IRO	10.25273952	31/IRO	4.60966169
14/OPI	-7.44652246	31/OPI	0.98451165
15/FILLER	8.80172209	32/FILLER	3.48689513
15/IRF	-10.04905814	32/IRF	11.37692758
15/IRO	-3.01200217	32/IRO	-9.73424033
15/OPI	0.39990375	32/OPI	7.75177753
16/FILLER	-4.76875661	33/FILLER	-6.03673236
16/IRF	-1.10095774	33/IRF	3.26071129
16/IRO	5.93609823	33/IRO	10.29776725
16/OPI	2.31094818	33/OPI	6.67261721
17/FILLER	-11.59448329		
17/IRF	7.55483871		
17/IRO	14.59189468		
17/OPI	-10.14442327		

## Appendix M. Descriptive Statistics for the learners' response times for the full dataset (including correct and incorrect answers)

**Table 30**

*Learners' Response Times*

	Mean	SD	Min.	Max.
Indirect refusals	8.95	2.74	4.15	14.58
Irony	10.04	3.19	3.61	19.57
Indirect opinions	9.04	2.19	4.36	12.99
Fillers/Literal	8.56	2.65	3.54	15.48
Overall	9.15	2.69	3.91	15.65

## Appendix N. Learner response time Accuracy Model - Random effects

**Table 31**

*Response Time Model random effects - Participant Intercept*

Participant	Random effect size	Participant	Random effect size
1	3.07059842	18	-0.22770406
2	2.75553930	19	-3.43363188
3	1.79699376	20	-2.13270307
4	5.90102995	21	-2.18634595
5	-1.09217348	22	1.03631821
6	-1.87682823	23	0.03890555
7	1.47856649	24	2.33916118
8	1.31424263	25	-1.55515269
9	-4.85756964	26	1.94496025
10	-2.85289510	27	-0.24790521
11	3.14724216	28	-1.64141363
12	3.68771971	29	-0.43306055
13	0.57968473	30	-1.15994247
14	0.85498732	31	0.19135775
15	-1.20850026	32	-1.70438369
16	-1.43744539	33	-1.45234490
17	-0.63730721		

**Table 32**

*Response Time Model random effects - Participant/Condition Intercept*

Participant/ Condition	Random effect size	Participant/ Condition	Random effect size
1/FILLER	0.876260791	18/FILLER	-0.396391032
1/IRF	0.101792339	18/IRF	-0.908845618
1/IRO	-0.156677545	18/IRO	0.775149150
1/OPI	-0.072928852	18/OPI	0.474585500
2/FILLER	2.449019547	19/FILLER	-0.254910755
2/IRF	-0.527804367	19/IRF	-0.269757847
2/IRO	0.802478882	19/IRO	-0.455558952



2/OPI	-2.052041796	19/OPI	0.143292789
3/FILLER	-1.156128623	20/FILLER	0.280155183
3/IRF	1.444707186	20/IRF	-0.678717671
3/IRO	1.345138824	20/IRO	0.232131081
3/OPI	-1.195706967	20/OPI	-0.353406886
4/FILLER	0.796660740	21/FILLER	-0.488121263
4/IRF	-0.211697650	21/IRF	0.282601464
4/IRO	2.851889971	21/IRO	-0.600990291
4/OPI	-1.998499367	21/OPI	0.273596547
5/FILLER	0.015143145	22/FILLER	-0.185675633
5/IRF	-0.132512858	22/IRF	0.187328048
5/IRO	0.353200601	22/IRO	-0.673460518
5/OPI	-0.502044035	22/OPI	0.924406741
6/FILLER	1.439992272	23/FILLER	-0.578116402
6/IRF	-0.413237332	23/IRF	-0.374432945
6/IRO	-1.302923042	23/IRO	-0.387965139
6/OPI	-0.181301669	23/OPI	1.349997567
7/FILLER	-0.761847929	24/FILLER	-1.673342212
7/IRF	0.363136735	24/IRF	1.161567006
7/IRO	-0.762706124	24/IRO	1.694000787
7/OPI	1.521812298	24/OPI	-0.612063904
8/FILLER	-0.244511330	25/FILLER	-0.137890455
8/IRF	2.225061986	25/IRF	-0.059794898
8/IRO	-0.410771268	25/IRO	-0.425859134
8/OPI	-1.249437727	25/OPI	0.244481911
9/FILLER	-0.132608801	26/FILLER	0.481717026
9/IRF	0.046882310	26/IRF	-0.826214970
9/IRO	-1.239107077	26/IRO	-0.065487203
9/OPI	0.140819370	26/OPI	0.884061839
10/FILLER	0.186262689	27/FILLER	0.387017322
10/IRF	-0.793565508	27/IRF	-0.286448737
10/IRO	-0.458038561	27/IRO	-0.286023808
10/OPI	0.369959006	27/OPI	0.125029266
11/FILLER	0.391383450	28/FILLER	-0.291567475
11/IRF	1.620843411	28/IRF	-0.474961577
11/IRO	-0.515023409	28/IRO	-0.452978462
11/OPI	-0.730075099	28/OPI	0.819419161
12/FILLER	-0.375100534	29/FILLER	-0.122079308
12/IRF	0.039381360	29/IRF	-1.381534871
12/IRO	1.029320474	29/IRO	1.567643658
12/OPI	0.205266408	29/OPI	-0.169586346
13/FILLER	0.643207253	30/FILLER	0.931413754
13/IRF	-0.270727594	30/IRF	-0.272593368
13/IRO	-1.246908154	30/IRO	0.514213906
13/OPI	1.015724452	30/OPI	-1.455765874
14/FILLER	0.865339712	31/FILLER	-0.800291951
14/IRF	-0.056432093	31/IRF	-1.715262116
14/IRO	-0.759826243	31/IRO	2.403693729
14/OPI	0.159318546	31/OPI	0.158503065
15/FILLER	-0.107029428	32/FILLER	-0.522689530
15/IRF	1.347919984	32/IRF	-0.687749345

15/IRO	-0.697061713	32/IRO	-0.030029081
15/OPI	-0.838396204	32/OPI	0.825030889
16/FILLER	-1.147762886	33/FILLER	0.430060059
16/IRF	0.604348757	33/IRF	0.748952036
16/IRO	-0.259912709	33/IRO	-1.527534437
16/OPI	0.452954970	33/OPI	-0.005481224
17/FILLER	-0.797567395		
17/IRF	0.167768744		
17/IRO	-0.854018192		
17/OPI	1.328475625		

## Appendix O. Difference between high and low-scoring mean for follow-up interview item selection

Table 33

*Difference between high and low-scoring mean for follow-up interview item selection*

	T1_LastClick	Q1 score	T2_LastClick	Q2 score	T3_LastClick	Q3 score	T4_LastClick	Q4 score	T5_LastClick	Q5 score	T6_LastClick	Q6 score								
Mean low-scoring group	8.203	1.000	8.816	0.182	9.105	1.000	9.248	0.818	12.281	0.727	16.054	0.364								
Mean high-scoring group	8.186	1.000	6.241	1.000	6.924	1.000	7.013	0.909	9.766	0.818	7.079	1.000								
Difference High-Low Mean	-0.017	0.000	-2.575	0.818	-2.181	0.000	-2.235	0.091	-2.515	0.091	-8.975	0.636								
	T7_LastClick	Q7 score	T8_LastClick	Q8 score	T9_LastClick	Q9 score	T10_LastClick	Q10 score	T11_LastClick	Q11 score	T12_LastClick	Q12 score	T13_LastClick	Q13 score	T14_LastClick	Q14 score	T15_LastClick	Q15 score	T16_LastClick	Q16 score
10.088	0.364	8.316	0.545	13.106	0.000	5.926	1.000	13.210	0.455	7.788	0.455	10.178	6.909	0.576						
7.370	0.727	5.636	1.000	13.293	0.545	5.524	1.000	9.883	1.000	10.680	0.818	8.133	10.818	0.902						
-2.718	0.364	-2.680	0.455	0.187	0.545	-0.403	0.000	-3.327	0.545	2.892	0.364	-2.046	3.909	0.326						
	T17_LastClick	Q17 score	T20_LastClick	Q20 score	T21_LastClick	Q21 score	T23_LastClick	Q23 score	T24_LastClick	Q24 score	T25_LastClick	Q25 score	T26_LastClick	Q26 score						
6.551	0.818	10.499	0.455	10.334	0.636	11.849	0.364	11.814	0.000	17.615	0.818	10.645	0.182							
7.256	0.818	7.869	0.545	8.077	1.000	13.553	0.545	13.019	0.545	11.431	0.909	10.239	0.364							
0.705	0.000	-2.630	0.091	-2.257	0.364	1.704	0.182	1.205	0.545	-6.185	0.091	-0.406	0.182							
	T27_LastClick	Q27 score	T28_LastClick	Q28 score	T29_LastClick	Q29 score	T30_LastClick	Q30 score	T32_LastClick	Q32 score	T39_LastClick	Q39 score	T40_LastClick	Q40 score						
8.570	1.000	11.299	0.818	11.097	0.091	11.503	0.545	8.985	0.818	10.897	6.545	8.727	0.727							
8.432	0.909	10.468	1.000	11.857	0.273	10.126	0.909	8.301	0.909	10.052	8.727	2.182	0.182							
-0.138	-0.091	-0.831	0.182	0.760	0.182	-1.378	0.364	-0.684	0.091	-0.845	2.182	0.636								
	T33_LastClick	Q33 score	T34_LastClick	Q34 score	T36_LastClick	Q36 score	T37_LastClick	Q37 score	T38_LastClick	Q38 score	T39_LastClick	Q39 score	T40_LastClick	Q40 score						
9.252	0.909	12.118	0.091	7.268	0.455	8.293	0.364	6.981	0.818	10.370	0.273	8.615	0.364							
6.646	1.000	12.465	0.727	7.986	1.000	11.872	0.909	5.938	0.909	13.685	1.000	7.529	1.000							
-2.607	0.091	0.347	0.636	0.718	0.545	3.579	0.545	-1.043	0.091	3.315	0.727	-1.085	0.636							
	T41_LastClick	Q41 score	T42_LastClick	Q42 score	T44_LastClick	Q44 score	T46_LastClick	Q46 score	T48_LastClick	Q48 score	T54_LastClick	Q54 score	T55_LastClick	Q55 score						
6.541	1.000	8.815	1.000	8.767	1.000	11.686	0.818	11.080	0.545	9.149	7.636	0.636								
6.349	0.909	7.817	0.909	8.497	1.000	8.702	1.000	9.959	0.909	8.954	11.273	0.939								
-0.192	-0.091	-0.998	-0.091	-0.270	0.000	-2.985	0.182	-1.121	0.364	-0.195	3.636	0.303								
	T49_LastClick	Q49 score	T50_LastClick	Q50 score	T51_LastClick	Q51 score	T52_LastClick	Q52 score	T53_LastClick	Q53 score	T54_LastClick	Q54 score	T55_LastClick	Q55 score						
8.470	1.000	10.745	0.909	7.794	1.000	10.340	0.455	10.681	0.727	10.713	0.364	15.191	0.545							
6.868	0.909	7.892	1.000	6.558	1.000	11.306	0.545	9.736	0.909	10.606	0.818	10.100	1.000							
-1.602	-0.091	-2.853	0.091	-1.236	0.000	0.966	0.091	-0.946	0.182	-0.107	0.455	-5.091	0.455							
	T56_LastClick	Q56 score	T57_LastClick	Q57 score	T58_LastClick	Q58 score	Average time for fillers			Total score										
6.646	0.909	6.868	0.909	10.614	0.909	9.806	7.727	0.773	28.818											
4.547	1.000	6.469	1.000	5.995	1.000	8.008	9.182	0.918	40.818											
-2.099	0.091	-0.399	0.091	-4.619	0.091	-1.799	1.455	0.145	12.000											