Ratings of humor produced by L2 and heritage speakers of Spanish

[Classificazione dell'umorismo prodotto da parlanti di spagnolo L2 e di spagnolo come lingua etnica]

Rachel L. Shively [1]

[1]*Illinois State University*
E-mail: rshivel@ilstu.edu

ABSTRACT

IT Questo studio esamina la classificazione dei tentativi di umorismo da parte di parlanti di spagnolo come seconda lingua (L2) e come lingua etnica. È stato chiesto a 31 parlanti di spagnolo come L2 e come lingua etnica che studiavano negli USA, di svolgere un esercizio di completamento di un testo e di fornire un commento umoristico in risposta a quattro spunti. Successivamente, i tentativi di umorismosono stati valutati in termini della loro effettiva carica umoristica. I risultati indicano che per l'intero gruppo le valutazioni in termini di umorismo sono state basse e che non ci sono state differenze significative nella valutazione dei due gruppi bilingui. L'analisi mette in evidenza quali siano stati gli aspetti nei tentativi di umorismo che gli osservatori hanno indicato come effettivamente divertenti.

Parole chiave: umorismo nella seconda lingua, umorismo nella lingua etnica, valutazione dell'umorismo, esercizio di completamento

EN This study examines the ratings of attempts at humor by second language (L2) and heritage speakers of Spanish. Thirty-one L2 and heritage speakers studying Spanish in the United States were asked to provide a humorous comment in response to four prompts. Participants' humor attempts were then rated in terms of their funniness by four English-Spanish bilingual speakers. The goal of this article is to analyze to what extent L2 and heritage speakers' attempts at humor were rated as funny and if there were differences between the two groups, as well as what reasons raters provided for judging humor as successful or unsuccessful. Findings indicated that funniness ratings were low for the group as a whole and, further, that there was no significant difference in ratings between the two bilingual groups. The analysis highlights aspects of participants' humor attempts that raters pointed to in their judgements of success in being amusing.

Key words: second language humor, heritage language humor, humor rating, discourse completion task

1. Introduction

The goal of the present study is to examine ratings of attempts at humor which were produced by two different groups of bilingual speakers: second language (L2) speakers and heritage speakers. The linguistic profile of each group is often distinct, due to different experiences with the language in question. Many L2 speakers develop their L2 proficiency through coursework and formal study, while heritage speakers, that is, bilingual individuals who grew up speaking a minority/minoritized language, typically gain skills in their heritage language through using the language with family and in bilingual communities (e.g., Carreira & Hitchins Chik, 2018; Potowski, 2018). While research exists on both L2 and heritage speakers' use of humor, the former has received considerably more attention than the latter.

In the case of L2 humor, speakers at all levels of language proficiency can and do use humor successfully to amuse others, as well as to achieve additional interactional goals (e.g., Bell & Pomerantz, 2015). However, previous research suggests that producing humor in an L2 can be challenging. In a fast-paced conversation, for example, L2 users have described not being able to think of a humorous reply quickly enough to insert it into the ongoing interaction (Bell & Attardo, 2010). Another challenge is that native or expert speakers may not perceive L2 speakers as legitimate users of humor, may not recognize their attempts at humor, or may perceive an utterance intended as humor as language error (Bell, 2006; Chiaro, 2009). Lack of shared cultural knowledge or cues to the humorous intent can further result in L2 users' humor not being understood in intercultural interactions (Shively, 2018). When humor fails to amuse in an interaction, that failure may be evident through the reaction of the interlocutor (e.g., absence of laughter) but it may not be clear why the attempt at humor was unsuccessful (Shively, 2018). Although research on humor in an L2 has examined L2 speakers' reports of both success and failure with humor, as well as analyzed successful and failed L2 humor in naturalistic interactions, few previous studies have systematically collected data on whether L2 speakers' attempts at humor are successful in being amusing. Further, this issue has not been addressed, to my knowledge, with heritage speakers.

With these gaps in the literature in mind, the present study focuses on humor produced by a group of L2 and heritage speakers who were studying Spanish in the United States. Using a discourse completion task, participants were asked to provide a humorous comment in response to four different prompts. Participants' attempts at humor were then rated in terms of their funniness by four bilingual speakers of English and Spanish. This article examines whether L2 and heritage speakers' attempts at humor were rated as funny or not and if there were differences between the two groups, as well as what reasons raters provided for judging humor as successful or unsuccessful in being amusing. The article concludes with suggestions for how these data can inform language pedagogy.

2. Literature review

Research has demonstrated bilingual speakers' linguistic creativity (i.e., using language in original and imaginative ways) as they communicate in a humorous mode (e.g., Bell, 2011; Belz, 2002; Forman, 2011; Martínez & Morales, 2014). This section begins with a discussion of research on humor in an L2, focusing on approaches to determining success in being amusing with humor in an L2, and ends with a discussion of humor in a heritage language.

2.1 Humor in an L2

A growing body of literature on L2 humor suggests that it is commonplace for L2 speakers to adopt a humorous mode of communication—where the intent of the speaker is to amuse a hearer—in interactions with expert speakers of the L2 and with other bilinguals (for an overview see Bell & Pomerantz, 2019). In languages such as English, German, and Spanish, L2 speakers have been observed to produce humor in a variety of settings, including informal conversations with friends, interactions with peers and instructors in the language classroom, written discourse in computer-mediated chat rooms, and jocular exchanges in business meetings (e.g., Adelswärd & Öberg, 1998; Ahn, 2016; Cheng, 2003; Vandergriff & Fuchs, 2012). These studies have highlighted how L2 speakers at all proficiency levels, from beginning to advanced, creatively marshal the linguistic and non-verbal resources in their L2 repertoires to amuse their interlocutors, as well as carry out other social functions such as build rapport and negotiate identities. While humor can create social bonds, some studies have also revealed the potential of humor to create conflict and threaten another's face, particularly when engaging in playful teasing, that is, poking fun at one's interlocutor (e.g., Shively, 2018). Previous research further suggests that L2 speakers are often successful in using humor in their L2 to amuse others (e.g., Bell, Skalicky, & Salsbury, 2014; Davies, 2003; Shively, 2018), but there are also cases when L2 speakers' humor has been observed to fail because some aspect of the humorous utterance (e.g., linguistic form, pragmatic force, meta-message) was not understood or appreciated by the interlocutor (e.g., Bell, 2015; Bell & Attardo, 2010). For instance, Shively (2018) described instances in which Spanish speakers did not understand English-Spanish bilingual word play produced by L2 speakers.

The vast majority of studies that have examined instances of humor in an L2 have been based on the collection and analysis of naturalistic discourse data. In such studies, the humor produced by L2 speakers emerges spontaneously as the interaction unfolds. Determining the success or failure of a humorous utterance in naturalistic data typically involves careful analysis of the discourse context, including the reactions of the speaker and hearer in subsequent turns, for example, whether the hearer produces laughter or other humor support in response. Interviews with the parties to the talk to obtain their perspectives on whether humor was successful or not have also been used in previous studies to dig deeper into perceptions of humor use (e.g., Bell, 2006; Shively, 2018). However, in the absence of introspective data or of metalinguistic comments in the talk about the humor (e.g., "that joke wasn't funny because…"), it can be difficult for both the L2 humorist and the humor researcher to determine why a particular instance of humor was or was not funny for the hearer (Shively, 2018).

Few previous L2 humor studies have engaged raters to systematically evaluate the funniness of attempts at humor. To my knowledge, only two previous studies have employed ratings to assess L2 humor samples and both involved data elicitation instruments rather than naturalistic recordings (Petkova, 2013; Shardakova, 2013). In the first of these, Petkova asked two native-English-speaking raters to evaluate humor samples collected using closed role plays (i.e., oral responses to a prompt) in which intermediate and advanced L2-English-speaking participants were asked to provide a humorous remark orally in English in response to three scenarios. Raters evaluated each attempt at humor on a five-point scale from "very funny" to "not funny at all," with an additional category for "unintelligible." Prior to scoring, the raters

participated in a norming session to establish inter-rater reliability. It does not appear that raters were asked to provide an explanation for their scores. Using the resulting ratings, Petkova analyzed whether L2 speakers' humor was perceived by raters as funnier after an eight-week pedagogical intervention to teach about humor in English—the answer was no. However, the author reported that a large percentage of the L2 humor produced, both pre- and post-intervention, was perceived by raters as "very funny" or "funny." At the same time, although participants possessed intermediate or advanced proficiency in English, a fairly high percentage of respondents (19-38%, depending on the scenario) decided not to attempt a humorous response.

A second study, by Shardakova (2013), in which humor ratings were employed, involved L2 speakers of Russian. In this case, data were collected through a written discourse completion task (DCT), a questionnaire consisting of a series of scenarios to each of which respondents provided a single-turn response. The DCT was designed to elicit speech acts in Russian (apologies, compliments, invitations, requests), but some participants employed humor in their responses even though the data collection measure did not prompt respondents to do so. Focusing just on those DCT responses that incidentally contained attempts at humor, Shardakova asked native speakers of Russian to evaluate L2 speakers' humorous replies. In contrast to the largely positive reception to L2 English humor in Petkova's (2013) study, attempts at humor in L2 Russian in Shardakova's study were less successful: only about one third of the L2 Russian humor was rated by Russian native speakers as "funny" or "somewhat funny."

Asking native or expert speakers to assess L2 speakers' performance through rating, as was done in Petkova (2013) and Shardakova (2013), is commonplace in the L2 pragmatics literature. One aspect that has received little attention in any area of L2 pragmatics, however, is variability in raters' perceptions of the categories that they are asked to use to score L2 language samples. When examining precisely this issue, Taguchi (2011) observed that four native-English-speaking raters diverged in the reasons that they gave for scoring the pragmatic appropriateness of L2 English speakers' speech act performance with requests and opinions. By interviewing the raters and examining their discourse during evaluation sessions, Taguchi reported that while some raters oriented more to linguistic forms, others focused on politeness strategies or situation-specific criteria. This research on speech act rating demonstrates that rater perception and interpretation can vary even when evaluating a speech act such as a request, which is often has routine realizations in specific settings (e.g., service encounters) in a given speech community. In the case of assessing funniness, we might expect even greater variability among raters, given that, first, humor often involves an attempt at being creative with language—bending the rules rather than adhering to them—and, second, that sense of humor varies from person to person (e.g., Bell, 2015; Kuipers, 2015; Ruch, 2008).

2.2 Humor in a heritage language

Although heritage speakers all share an ancestral connection to their heritage language, they are a heterogeneous group in terms of their skills in that language. Factors such as generation of immigration, language of schooling, family language practices, and motivation, attitudes, and identity influence individuals' heritage language proficiency (e.g., Beaudrie & Fairclough, 2012). However, in comparison to L2 speakers, heritage speakers as a group tend to have stronger aural skills, greater ease with spontaneous interaction, and deeper cultural knowledge related to their ancestry (e.g., Carreira & Hitchins Chik, 2018;

Chang, 2017). For some heritage speakers, the heritage language is not their dominant or preferred language, while for others it is or both languages are viewed equally (e.g., Vaid, 2006).

The existing research on heritage speakers' use of humor has primarily documented the characteristics and social functions of that humor, rather than focusing on its success or failure. Most studies have considered heritage languages in the context of the United States, where non-English languages such as Spanish are not only minority languages—in the sense that less than 50% of the population speaks them and they have no official or legal status—but are also minoritized, that is, discriminated against and deemed inferior (Potowski, 2018). In this context, several authors have highlighted how humor in the heritage language may serve as a means to resist the marginalization of that language and its speakers (e.g., Leung, 2014; Zentella, 2003). Zentella (p. 63) observed, for instance, that "Latinos have a good deal of fun at the expense of gringos [white Americans], and language play is at the heart of their defense against their marginalization, exploitation and stigmatization."

Some research has demonstrated heritage speakers' considerable creativity and skill with humor in their heritage language. In a study on humor produced by adolescent Spanish heritage speakers in the United States, Martínez and Morales (2014) observed instances in which individuals employed sophisticated humorous utterances involving double entendre in Spanish. These authors also reported that for heritage speakers with lower proficiency in Spanish, using humor in Spanish was a way for them to position themselves as competent speakers of the language. Further, bilingual speakers—L2 or heritage—may specifically draw on linguistic resources from their entire bilingual repertoire and produce humor involving code-switching and bilingual puns (e.g., Martínez & Morales, 2014; Vaid, 2006). Some authors have also pointed to research on bilingual humor practices to highlight the considerable linguistic competence of heritage speakers, which may be overlooked in contexts in which there is a deficit view of bilingualism and discrimination against heritage languages and their speakers (e.g., Leung, 2014; Martínez & Morales, 2014).

One previous study spotlighted differences in self-reported humor preferences based on linguistic and cultural background. Using surveys with Mexican-American and Anglo-American English-Spanish bilinguals, Vaid (2006) observed that the two groups shared perceptions about some uses of humor, but differed in certain ways. Mexican-Americans reported valuing being creative with language more than Anglo-Americans, while the latter placed more importance on being able to use self-deprecating humor. Looking specifically at the Mexican-American participants' choice of language to be funny, the researcher found no preference for either English or Spanish in telling humorous personal narratives or playfully teasing, but did find a preference in some participants for using English to be sarcastic (i.e., aggressive, disparaging humor; see e.g., Dynel, 2018) or to express hostility. The findings also suggested that the generation of immigration of Mexican-Americans influenced some aspects of humor. For instance, those with parents born in Mexico were more likely to say that their sense of humor changed based on the language of interaction, whereas few of those whose parents were born in the United States experienced changes in their sense of humor in English and Spanish.

2.3 Aims of this study

As this review of previous work indicates, the systematic analysis of the success of L2 and heritage speakers in amusing others through their attempts at humor has received little attention. Such an approach can expand upon existing research and provide greater insight into why bilingual speakers' humor amuses or fails to amuse. Hence, this study was designed to address this underresearched area and the following questions were developed to guide the analysis presented in the following sections:

- Research question 1: To what extent were attempts at humor produced by L2 and heritage speakers rated as funny? Were there any differences between L2 and heritage speakers?
- Research question 2: What features did raters point to as contributing to successful and unsuccessful
   humor? To what extent was there divergence?

3. Methods

3.1 Participants

The participants in this study were 31 undergraduate students at a mid-sized public university in the United States. All spoke both English and Spanish and were pursuing an academic minor or major in Spanish. They were also all enrolled in one of two sections of a Spanish-language upper-division course, taught by the same instructor. As shown in Table 1, 10 of the participants indicated that they had spoken Spanish with their family growing up (heritage speakers), while the other 21 reported no contact with Spanish in childhood outside of formal coursework (L2 speakers). The ancestral homelands of heritage speakers included Mexico and several countries in Central America. Heritage speakers, on average, had spent less time formally studying Spanish compared to L2 speakers. Twenty-two of the students identified as female, and nine as male. While most students had not completed university studies outside of the United States, five students indicated that they had studied abroad in short-term or semester-long programs. Although Spanish proficiency was not measured in this study, all students had successfully completed prior university coursework designed for those at the American Council on the Teaching of Foreign Languages intermediate level or higher (A2 or higher in the Common European Framework of Reference for Language framework).

Table 1. *Participant demographics*

	LANGUAGE BACKGROUND	GENDER		STUDIED ABROAD	MEAN YEARS FORMAL STUDY OF SPANISH
		Female	*Male*		
L2 SPEAKERS	21 (68%)	16 (76%)	5 (24%)	5 (24%)	6.57
HERITAGE SPEAKERS	10 (32%)	6 (60%)	4 (40%)	0 (0%)	5.00
TOTAL	31 (100%)	22 (71%)	9 (29%)	5 (16%)	5.78

3.2 Data collection

3.2.1 Data collection instrument

Humor samples were collected by means of a written discourse completion task (DCT). A written DCT is a production questionnaire consisting of a series of scenarios to each of which respondents typically provide a single-turn response (e.g., Ogiermann, 2018). Although written DCTs have been employed for decades in L2 pragmatics research, primarily to collect speech act data, it is only recently that an attempt to collect L2 humor data with DCTs was made. Bell, Shardakova, and Shively (2021) explored the use of a DCT to collect L2 humor samples and the results indicated that this DCT was, indeed, effective in achieving this aim.

The primary advantage of a written DCT as a data collection instrument—compared to role plays (i.e., simulated spoken interactions) or recordings of naturalistic interactions, for example—lies in its practicality: a large corpus of data can be collected and analyzed quickly and easily; data are comparable over time and across participants; data can be collected in situations that would be difficult for researchers to access in real life; and the questionnaire can be translated into any language, allowing cross-linguistic and cross-cultural comparisons. Bell et al. (2021) also argued that a potential benefit of written DCTs for collecting L2 humor data specifically is the fact that the written format provides sufficient time for L2 speakers to create and edit their humorous responses—an issue that L2 users have reported struggling with, as described above. Further, given that most previous studies concerning L2 humor have examined instances of humor in naturalistic conversation, what is currently known about L2 humor is primarily how it is used spontaneously during an ongoing spoken interaction, not what L2 speakers can do with humor given more planning and editing time.

On the other hand, the written DCT format also has disadvantages, such as not going beyond a single turn and not permitting the respondent to provide prosodic and non-verbal information. Previous research on DCTs also suggests that written responses of what participants think they would say in a given scenario may differ from what they would actually say in a similar situation in real life (e.g., Golato, 2003). Further, asking participants to write what they think they would *say* represents an indirect measure of oral speech. In

this study, the choice was made to use a written DCT due to the benefits listed above, although its limitations were recognized.

A DCT was designed with four items (shown below), each of which included a short description of the context, an utterance, and a prompt for the test-taker to respond to the utterance with humorous irony or sarcasm. The focus on irony is due to the fact that the data were collected as part of a pedagogical intervention designed to teach humorous irony (see Shively, Acevedo, Cano, & Etxeberria-Ortego, 2022, for information about the pedagogical intervention). Using Dynel's (2014, 2018) approach, irony was defined in this study as having two key features that differentiate it from other figures of speech and types of humor: overt untruthfulness (i.e., the speaker pretends to mean what is said) and evaluative implicature (i.e., the speaker expresses a judgment). Further, although irony and sarcasm can be considered separate phenomena, the DCT instructions included both terms for clarity, since in common parlance in American English the two are often used interchangeably (e.g., Dynel, 2017).

The utterances in Situations 1 ("Green pants") and 3 ("Paris photo") below were taken verbatim from two language corpora of naturalistic audio-recordings in Spanish; the former is from the Val.Es.Co corpus of spoken Peninsular Spanish (https://www.uv.es/corpusvalesco/) and the latter from data published in Shively (2018) involving an L2 speaker interacting with a Peninsular Spanish speaker. In both corpora, the utterance included in the Situations 1 and 3 was followed by humorous irony. In contrast, Situations 2 ("Play soccer") and 4 ("Water the plants") were modified from scenarios with irony included in didactic materials in Spanish published as part of a research project (Yanguas Santos, 2006). The first scenario prompts the respondent to joke about a person that is not present in the interaction, while the remaining three scenarios elicit humor that will most likely target the interlocutor. Prior to using this DCT, the items were pilot tested with both L2 and expert speakers of Spanish to ensure the instructions and scenarios were clear. Note that although English translations of the prompting utterances are provided below, they were not provided to participants in the research.

DCT

Instructions: Imagine that you are studying abroad in a Spanish-speaking country. Below you will be presented with situations from everyday life in study abroad. For each question, first read the description of the situation and then the comment made. Try to think of a funny sarcastic or ironic comment that you could make in response and write it in the space below.

Situation 1
Your language partner and you have become good friends. One day you're sitting and chatting in Spanish. She's showing you some photos from a family event in which her father was wearing some very bright green pants, which she laughs about. Read your language partner's comment below and then write something humorously sarcastic or ironic in response.

Language partner: *Mira mi padre con esos pantalones, ¿eh?* ((laughs))
　　　　　　　　'Look at my father with those pants, eh?'
You say:

Situation 2

You get together with others to play soccer for fun on a regular basis in Granada. You've become good friends with one of the other people who plays, Sergio. There's a game this afternoon and you arrive a bit early and see Sergio sitting in his soccer gear smoking a cigarette. Read Sergio's comment below and then write something humorously sarcastic or ironic in response.

Sergio: *Vamos a jugar ya. Me encanta hacer deporte.*
            'Let's play. I love sports.'

You say:

Situation 3

You and your friend are chatting about a trip that she took to Paris. She mentions that, unfortunately, her photos from the trip were all accidentally deleted from her phone, except for one, which shows her in a McDonald's restaurant with her back turned. Read your friend's comment below and then write something humorously sarcastic or ironic in response.

Friend: *No tengo fotos de París.* ((laughs))
Solo tengo una que no sé quién hizo en blanco y negro
Que salgo yo dentro de un McDonald's, hablando por teléfono de
espaldas. ((laughs))
'I don't have photos from Paris. ((laughs))
I only have one that I don't know who took it in black and white
that I appear inside a McDonald's, on the phone,
with my back to the camera.' ((laughs))

You say:

Situation 4

You come home from class on a warm day and find your host mother out on the patio looking at some plants that are completely brown and withered. Read your host mother's comment below and then write something humorously sarcastic or ironic in response.

Host mother: *Creo que por fin ha llegado el momento de regar las plantas.*
((laughs))
'I think the time has finally come to water the plants.'

You say:

Finally, in addition to the DCT, participants completed a background questionnaire that served to collect demographic data. Participants were asked to provide age, gender, years formally studying Spanish, any previously study abroad experiences, and whether they spoke Spanish at home with their family.

3.2.2 Data collection procedures

Data included in the present analysis were collected as part of a larger project. As described in Shively et al. (2022), the overarching research involved a short-term pedagogical intervention to teach humorous irony in Spanish in a course that included a mixed group of L2 and heritage speakers. The instruction lasted for 2.5 hours and involved input and output practice with authentic samples of humorous irony in Spanish. The goal was to determine the effectiveness of instruction for both L2 and heritage speakers. For that reason, the design of the larger study included collection of humorous irony comprehension and production data prior to and after instruction (pre-test/post-test design). Consequently, the DCT and background questionnaire were administered at Time 1 (pre-test), three weeks prior to the instruction on humorous irony, and then the DCT was administered again at Time 2 (post-test), which occurred one week after instruction. At each administration, participants were given all the time they needed to complete the DCT and were told to leave an item blank if they could not think of something funny to say. The pre- and post-test DCT items (shown above) were the same during the pre- and post-tests, but they were presented to the participants in a different order at each time. In the present study, only the DCT data are examined and the focus is not on the pedagogical intervention. Further, data from both the pre- and post-test DCT are included in the analysis. Notably, Shively et al. (2022) found that there was no statistical difference in funniness ratings of pre- and post-test responses to the DCT.

3.3 Data analysis

Each of the responses collected in the four scenarios on the DCT was evaluated by four raters. All raters were English-Spanish bilinguals who had advanced proficiency in their L2. None of the raters were heritage speakers. As Table 2 shows, American English was the first language (L1) of Rater 1, Colombian Spanish was the L1 of Raters 2 and 3, and Peninsular Spanish for Rater 4. The choice to use bilingual raters was based on, first, not wanting to advance a monolingual target and, second, the suggestion from previous research that some humor may only be enjoyed by people who share the same languages (Vaid, 2006). Three raters were female and one male. They were all a few years older than the study participants and they all had experience teaching Spanish as an L2. Rater 1 was born and raised in the United States, while Raters 2 and 3 were from Bogotá, Colombia, and Rater 4 had grown up in the Basque Country in northern Spain. All raters had been residing in the United States for at least two years at the time the ratings were conducted.

Table 2. *Raters' L1 and gender*

RATER	L1	GENDER
RATER 1	American English	Female
RATER 2	Colombian Spanish	Male
RATER 3	Colombian Spanish	Female
RATER 4	Peninsular Spanish	Female

The raters indicated the perceived funniness of each DCT response using the following three-point scale, similar to that employed in Petkova (2013): *not funny* (0), *somewhat funny* (1), and *very funny* (2). The category descriptions were also similar to those provided by Petkova. A response of *not funny* meant that the rater did not laugh or smile and was not amused by the response; the humor was unsuccessful. A rating of *somewhat funny* would be given in cases when the response at least made the rater smile and be amused to some extent. *Very funny* meant that the response made the rater laugh and think "That's a good one!". Although the researcher and raters discussed the category descriptions, there was no attempt prior to scoring to establish inter-rater reliability through a norming session, since it was assumed that sense of humor can vary among individuals and across cultures. Scoring was done individually and the ratings were blinded: raters were not told whether a response was from the pre-test or post-test or an L2 or heritage speaker. In addition to scoring each response, the raters were asked to briefly explain why they thought the response was funny or not.

The ratings were first analyzed quantitatively to determine the extent to which participants' humor was rated as funny. A Mann-Whitney U test was then used to compare ratings of humor attempts produced by L2 versus heritage speakers. The researcher subsequently compiled all the raters' scores and comments for each DCT scenario, carefully read through the data multiple times, and coded characteristics of the humor attempts that raters pointed to as reasons for their scores, as well as areas where ratings diverged.

4. Results

The goal of this section is to examine to what extent attempts at humor by L2 and heritage speakers were rated as funny, to determine whether ratings differed significantly between the two bilingual groups, and to describe the reasons raters provided for their funniness scores. To begin, an overview of the participant responses and a description of the funniness ratings is provided, which is followed by analysis of specific features in humor rated as funny or not funny.

4.1 Research question 1: Funniness ratings

All but one participant was able to produce a humorous response for each of the four DCT scenarios, once in the pre-test and a second time in the post-test. One participant was unable to think of a humorous reply for Situation 2 ("Play soccer") or 4 ("Water the plants") at either time and left the items blank. Hence, the data consist of 62 attempts at humor each for Situations 1 ("Green pants") and 3 ("Paris photo") and 60 attempts at humor each for Situations 2 and 4 for a total of 244 humor samples. All participant responses were evaluated by the four raters. In no case was a participant's humorous response to a particular scenario identical in the pre- and post-test. Although the DCT asked students to write a humorously ironic or sarcastic response, not all did use irony, as it was defined above (Dynel, 2018). Ironic responses were particularly less frequent in the pre-test before participants had received instruction about irony, but increased in the post-test (see Shively et al., 2022, for more details on the pedagogical intervention). No statistically significant differences were found between the L2 and heritage speakers at either the pre- or post-test with regard to funniness ratings, hence, the analysis will continue by looking at ratings for the group as a whole.

The mean funniness rating for all participants (L2 and heritage speakers) when the four raters' scores from 0-2 (0 = *not funny*; 1 = *somewhat funny*; 2 = *very funny*) are combined was .6058, which suggests that, in a global sense, the humor produced by participants in this study was not resoundingly successful. However, some participants' humor attempts were successful, whereas others were not and, similarly, some raters were amused by an answer that failed amuse the other raters. Indeed, there was considerable variation among the raters' mean funniness ratings. Rater 1 (L1 English, female) was the least amused by the humor with a mean score of .2823 (SD: .23267); Rater 2 (L1 Colombian Spanish, male) was the most amused with a mean of .7903 (SD: .34669); Rater 3 (L1 Colombian Spanish, female) was the second-least amused with .5565 (SD: .29197); and Rater 4 (L1 Peninsular Spanish, female) was the second-most amused at .7621 (SD: .37688). Notably, the L1-English-speaking rater was less amused than any of the three L1-Spanish-speaking raters.

All four raters agreed that a response was "somewhat funny" or "very funny" in only 17% of the cases (N = 41 humor samples). Twenty-eight different individuals produced these 41 humor samples, meaning that almost all participants produced at least one response that was viewed as humorous by all four raters. However, more than half (N = 23) of the humor samples that were funny to all raters were produced by 10 individuals (5 L2 speakers and 5 heritage speakers). Nine of those individuals produced successful humor in two items and three participants were successful with both pre- and post-test responses in one item. Therefore, about one third of the group was somewhat more successful with humor than the other two thirds. Looking just at the L1-Spanish-speaking raters, 27% (N = 67 humor samples) were either "somewhat funny" or "very funny" to all three. However, 81% (N = 198 humor samples) of the responses to the DCT scenarios were perceived as "somewhat funny" or "very funny" to at least one of the three L1-Spanish-speaking raters.

Funniness ratings were also observed to vary by item on the DCT, as shown in Table 3. For all four raters, responses to Situation 4 ("Water the plants") were the most successful, with 17 (28%) humor responses rated by all four raters as either "somewhat funny" or "funny." It seems, however, that the L1-English-speaking rater was more amused by humor attempts in Situation 4, compared to the other items because for all three L1-Spanish-speaking raters, the item with the most amusing responses was Situation 1 ("Green pants"), followed by Situation 4. Situation 2 ("Play soccer") was the item in which participants had the least success in producing a funny response in the view of the L1-Spanish-speaking raters.

Table 3. *Number and percent of responses rated as "somewhat funny" or "funny" by DCT situation*

SITUATION	ALL FOUR RATERS	ALL THREE L1-SPANISH-SPEAKING RATERS	AT LEAST ONE L1-SPANISH-SPEAKING RATER
SITUATION 1	8/62 (13%)	21/62 (34%)	55/62 (88%)
SITUATION 2	8/60 (13%)	12/60 (20%)	38/60 (63%)
SITUATION 3	8/62 (13%)	17/62 (27%)	54/62 (87%)
SITUATION 4	17/60 (28%)	17/60 (28%)	51/60 (85%)
TOTAL	41/244 (17%)	67/244 (27%)	198/244 (81%)

4.2 Research question 2: Reasons for funniness ratings

4.2.1 Ratings of successful humor

Although all raters provided an explanation for each of their funniness scores, it was clear that they were not always able to pinpoint exactly why they thought a humor attempt was or was not funny. Some rater comments did not go beyond "It's not funny" or "It really made me laugh." However, many of the raters' comments did highlight specific aspects of participants' humor that they thought made the utterance funny or not. Looking first at the reasons that raters provided for the responses they rated as "somewhat funny" or "very funny," the following features were mentioned multiple times by raters: exaggeration, similes and metaphors, self-deprecating humor, and written representations of prosodic, paralinguistic, and non-verbal information. Each of these aspects is explored below.

Exaggeration. Previous research has suggested that a common feature of humor in both L1 and L2 Spanish is exaggeration (Shively, 2018). A similar finding is true for this study, since many participant responses had some element of exaggeration and, in the funniness ratings, all four raters mentioned multiple times that a reason that they thought a response was funny was because something about it was exaggerated. For instance, in Situation 1 ("Green pants"), one heritage speaker wrote the following in (1):

(1) ¡Es muy duro no poder mirar! Puedo notar esos pantalones desde millas de distancia. ('It's very difficult not to see! I can notice those pants from thousands of miles away.')
[Heritage speaker, pre-test, Situation 1 ("Green pants")]

In explaining their scores for this response, Raters 1, 3, and 4 all noted that exaggeration made the response funny. Rater 4 wrote that "I like the exaggeration." Indeed, the response to the utterance in Situation 1 in the Val.Es.Co corpus (from which the utterance was taken) was *guapísimo* ('super handsome'), an exaggeration using the superlative. Similarly, one rater pointed to exaggeration in example (2) as making the comment funny:

(1) *¡Podemos usar 'Fotoshop' así que hay escargot en el menú! El McDonald's más elegante en el mundo de verdad.* ('We can use Photoshop so there's escargot on the menu! The truly most elegant McDonald's in the world.')
[L2 speaker, post-test, Situation 3 ("Paris photo")]

Similes and metaphors. On various occasions, three of the four raters highlighted similes and metaphors as contributing to the humor. All three L1-Spanish-speaking raters scored the response in (3) as "very funny" and remarked about the comparison with Power Rangers (i.e., a group of fictional television superheros, one of whom wears a bright green outfit). Rater 2 wrote "It's funny because of the connection with the character," while Rater 3 commented "Good simile" and Rater 4 replied "Hahaha best comparison!" Rater 1 did not find this response funny because she indicated that she was unfamiliar with the cultural reference.

(2) *Se parece como un Power Ranger.* ('He looks like a Power Ranger.')
[Heritage speaker, pre-test, Situation 1 ("Green pants")]

An instance in which a comparison was viewed as funny by all four raters is shown in (4), a response for Situation 2 ("Play soccer"). Here three raters highlighted the metaphor: Rater 1, "The image of the chimney smoking is funny"; Rater 2, "The comparison is funny"; and Rater 3, "Good metaphor."

(3) *No sabía que las chimeneas pueden jugar fútbol.* ('I didn't know that chimneys could play soccer.')
[Heritage speaker, post-test, Situation 2 ("Play soccer")]

Self-deprecating humor. Attempts at self-deprecating humor were met with a positive response from all raters. For instance, a few participants made fun of themselves in Situation 1 ("Green pants"), as in (5) below. Rater 4 was particularly enthusiastic about self-deprecating humor, as in her comment about her rating of "very funny" for (5): "I really like this answer! It's a good idea to mock yourself and be part of the joke."

(4) *Oy, ¡sí! ¡Tengo los mismos en mi armario!* ('Oh, yes! I have the same ones in my closet!')
[L2 speaker, pre-test, Situation 1 ("Green pants")]

Written representations of prosodic, paralinguistic, and non-verbal information. Although the DCT was a written questionnaire, in some cases participants indicated that their response would be delivered with prosodic, paralinguistic, or non-verbal information. Raters, for their part, occasionally

pointed to written representations of such features (e.g., laughter, eye rolling, winking) as part of the reason why the response was funny. For instance, in (6), one L2 speaker used a written representation of laughter:

(5) *Ja ja ja. Esos son muy impresiva. Él mira una lima.* ('Ha ha ha. Those are very impressive. He looks like a lime.')
[L2 speaker, pre-test, Situation 1 ("Green pants")]

In response, Rater 3 wrote that "The 'ja ja ja' [ha ha ha] makes me more willing to laugh, I think." Another L2 speaker indicated in (7), that s/he would roll her eyes, which has been discussed in the literature as one cue speakers use with irony in Spanish (Padilla García, 2008). Rater 3 highlighted the eye rolling in her explanation of the rating for (7), indicating that the response was "very funny" and commenting that "The eye rolling adds irony."

(6) *O sí, el cigarrillo lo dice todo. ((Poner los ojos en blanco))* ('Oh yes, the cigarrette tells it all. ((Eye roll))')
[L2 speaker, post-test, Situation 2 ("Play soccer")]

Multiple times Rater 2 mentioned that the delivery and intonation of the comment would need to be right to ensure it would be funny. In the case of example (8), Rater 2 gave the comment a rating of "very funny," but wrote the following caveat: "It sounds funny because of the irony, but you'd have to see if the intonation is right." Indeed, the participant may have included the ellipsis in (8) to indicate a certain intonation, such as trailing off or drawing out the sound.

(7) *¿Crees? ¿De veras? No me digas...* ('You think so? Really? You don't say...')
[Heritage speaker, post-test, Situation 4 ("Water the plants")]

In one case, shown in (9), Raters 1 and 3 highlighted sound lengthening with the word *poco* ('little') as a feature that influenced their rating of the response as funny. "The long 'poooco' adds a nice touch to the irony" was the comment of Rater 3, who thought (9) was "very funny."

(8) *No, no. Debes esperar un poooco más.* ('No, no. You should wait a liiittle longer.')
[L2 speaker, post-test, Situation 4 ("Water the plants")]

4.2.2 Ratings of unsuccessful humor

When humor was judged by raters to be unsuccessful, four different reasons appeared multiple times and in the explanations of more than one rater: they did not understand the meaning of the joke, they did not recognize a humorous frame (i.e., the response seemed to be in a serious key), they did not understand a reference, and they thought the response was too lengthy and involved.

Lack of understanding the humor. Although none of the raters, in their explanations of their scores, specifically mentioned grammatical or lexical errors in participants' responses, difficulties with expression in Spanish on some participants' part likely contributed to a lack of understanding of the

intended message in a few cases. All four raters had trouble understanding what the author of (10) meant in response to Situation 1 ("Green pants"), for instance:

(9) *¿Que puedes ser cómico a sus pantalones verdes posiblemente?* ('That you can be funny to your green pants possibly?')
[L2 speaker, post-test, Situation 1 ("Green pants")]

Humorous frame. At times participants' DCT responses did not seem to raters to have been produced with a humorous frame, but rather appeared to them as communicated seriously. Certainly, the instrument was limiting in this regard, since the humorous frame is often signaled via prosodic, paralinguistic, and non-verbal cues. Particularly in Situation 2 ("Play soccer"), the raters perceived many of participants' apparent attempts at playful teasing of the interlocutor in the scenario instead as serious communication and more specifically, as scolding or criticism. In (13), for example, all four raters indicated that the comment sounded like scolding, criticism, or rudeness.

(10)   *Tú también te encanta morir sus pulmones.* ('You also love your lungs dying.')
[L2 speaker, post-test, Situation 2 ("Play soccer")]

Unknown references. Humor was unsuccessful for all four raters at different times due to lack of shared knowledge with the humorist. As discussed previously, Rater 1 was unfamiliar with the reference to Power Rangers shown in (5). The participant in (12) was likely referencing a brand of vegetables common in the United States called Green Giant, whose mascot is an enormous man dressed all in green. Rater 2 wrote "I don't understand the reference" and Rater 3 asked "What green giant?" Rater 1, however, indicated that she understood the reference in (12) writing, "I understand that s/he's referencing the green giant from the vegetables, which is a funny comparison."

(11)   *¡Qué guapo! ¿Es posible que va de compras a la misma tienda que el gigante verde?* ('How handsome! Is it   possible   that   he   shops   at   the   same   store   as   the   green   giant?')
[L2 speaker, post-test, Situation 1 ("Green pants")]

Lengthy responses. Finally, several raters pointed out that participants at times seemed to be trying too hard, creating responses that were too long and complicated. About the example in (13), Rater 1 wrote "The part about seeing from space is funny, but the rest is too long." Rater 2 said "A lot of effort," and Rater 3 opined that it was "A little funny, but too long."

(12)   *Me sorprende que la cámara podía agarrar un buen foto con algo tan brillante en el cuarto, ¡se puede ver estos desde el espacio!* ('I'm surprised that the camera could get a photo with something so bright in the room, you   can   see   those   from   space!')
[Heritage speaker, post-test, Situation 1 ("Green pants")]

4.2.3 Divergence among raters

Although some similarities among raters were discerned, as described above, differences in the reasons that they gave for scoring an attempt at humor as successful or not were also evident. While some sources of divergence were response-specific, there were also three areas in which the raters diverged multiple times: calques, a lexical item, and teasing.

Calques. In a few instances, participants used literal translations from English to Spanish, like those in (14) and (15). The comment in (14) is a translation of the expression in English "to have a green thumb," meaning to be good with plants, but used ironically in this response to suggest that the person was not good with plants. This expression is not found in Spanish, as pointed out by Rater 3: "In Spanish we could say something like '*dejas morir hasta a un cactus*' [you even let a cactus die'] instead of associating it with green fingers." None of the L1-Spanish-speaking raters thought the expression was funny, even though they indicated knowing the expression in English. However, the L1-English-speaking rater did think the literal translation was funny, despite pointing out that it did not make sense in Spanish.

(13)   *¡Tienes   un   dedo   muy   verde!*   ('You   have   very   green   finger!')
[Heritage   speaker,   post-test,   Situation   4   ("Water   the   plants")]

(14)   *Al   menos   lo   estabas   amando.*   ('At   least   you   were   loving   it.')
[Heritage speaker, post-test, Situation 3 ("Paris photo")]

In contrast, neither Rater 1 nor Rater 2 found the literal translation of the McDonald's slogan "I'm loving it!" funny in (15). Neither indicated understanding the allusion to the line from the company's advertisements. However, the other two L1-Spanish-speaking raters rated (15) as "very funny." Rater 3 wrote that "Referring to McDonald's slogan—funny" and Rater 4 reported that the comment in (15) was "Hahahaha best pun."

Lexical item. A few participants in Situation 1 ("Green pants") compared the father wearing green pants to a clown or *payaso*, as in (16). Consistently, no matter the specific utterance, when making the comparison between father and clown, Rater 2 would say the humor sounded aggressive, but that it was funny. Rater 3 commented in all cases that the word *payaso* sounded offensive and not funny. Rater 4 took a more nuanced approach, indicating that the word *payaso* could only be used with someone you were close to: "Comparing the dad to somebody funny [like a clown] is smart, but you would say this to somebody you really know and feel close with, or else it sounds like an insult." Rater 1, on the other hand, did not think the *payaso* comparison was funny, but did not indicate that she thought the word was offensive or aggressive.

(15)   *¿Él   es   el   payaso   de   la   fiesta?*   ('Is   he   the   clown   of   the   party?')
[L2 speaker, pre-test, Situation 1 ("Green pants")

Teasing. As indicated previously, the context of Situations 2-4 led many respondents to poke fun at the interlocutor in the scenario, that is, to playfully tease him or her. In some cases, they also used teasing in Situation 1 ("Green pants"). Although raters were sometimes in agreement about the success or failure

of teasing to be funny, ratings of teasing were an area that saw considerable disagreement. In (17), one participant teased the interlocutor that the green pants were actually his/hers, not the father's. Everyone except Rater 3 thought the comment was "somewhat funny" or "very funny." Rater 3 explained her score of "not funny" by taking issue with making fun of the interlocutor's style.

(16)   *Pensé   que   esos   eran   tus   pantalones.*   ('I   thought   those   were   your   pants.') [L2 speaker, pre-test, Situation 1 ("Green pants")]

Similar disagreement among raters was observed with example (18). Raters 1 and 4 thought it was funny, but Rater 3 argued that it would be offensive and not funny, if the interlocutor were overweight. Rater 2 did not understand the joke in this response.

(17)   *¿Cuándo   no   te   encuentras   en   un   McDonald's?* ('When   are   you   not   in   a   McDonald's?') [Heritage speaker, pre-test, Situation 3 ("Paris photo")]

On the other hand, some teases were perceived as funny by all raters, as was the case with (19). In this case, the strongest blow of the tease falls on FIFA, the international soccer organization, rather than on the interlocutor who, in fact, may be viewed as being elevated playfully as being good enough to play in competitive soccer matches.

(18)   *Sigue fumando, el único deporte que vas a poder jugar será FIFA.* ('Keep smoking, the only sport you'll   be   able   to   play   will   be   FIFA.') [Heritage speaker, pre-test, Situation 2 ("Play soccer")]

5. Discussion

This study set out to examine whether attempts at humor produced by two groups of English-Spanish bilingual speakers—L2 and heritage speakers of Spanish—using a written DCT were viewed by four bilingual raters as funny or not. The results indicated no significant difference between L2 and heritage speakers with regard to the funniness ratings of their humor attempts. For the group as a whole, the ratings were low, indicating that the raters often did not find participants' responses to the scenarios amusing. The analysis also delved into the reasons raters provided for humor success or failure. As anticipated, raters varied considerably in many cases in their evaluation of a humor attempt as funny or not. Indeed, previous research indicates that individuals differ in the types of humor that they appreciate, in whether they find a particular stimulus funny, and in their propensity to laugh (e.g., Kuiper, 2015; Ruch, 2008).

A pronounced difference between the sole rater whose L1 was English and the other three raters whose L1 was Spanish was observed. The L1-English-speaking rater (Rater 1) was substantially less amused by participants' attempts at humor than the other three raters. Even though Rater 1 had the same nationality as the participants and shared with some the experience of learning Spanish as an L2, those characteristics held in common did not lead to greater appreciation of her co-nationals' humor, compared to the other raters.

For their part, the three L1-Spanish-speaking raters from Colombia and Spain were in agreement that a humor attempt was either "somewhat funny" or "very funny" in 27% of the cases. With their scores compiled as a group in this way, the three L1-Spanish-speaking raters in this study displayed similar perceptions to the native Russian speakers in Shardakova (2013), the latter of whom rated 28% of L2 Russian speakers' written humor as funny. These high rates of unsuccessful humor lend weight to reports from some L2 speakers that they struggle with using humor in their L2 (e.g., Bell & Attardo, 2010; Shively, 2018). At the same time, while the three L1-Spanish-speaking raters were in agreement less than a third of the time that a humor attempt was funny, the findings also indicated that 81% of the humorous responses produced by L2 and heritage speakers were funny to at least one of these three raters.

Just as funniness scores varied, so did raters' reasons for their funniness scores. Further, raters could not always pinpoint why a specific utterance seemed funny or not to them. However, an analysis of raters' explanations of their scores indicated that certain features appeared multiple times associated with successful and unsuccessful humor. Exaggeration was mentioned frequently by raters as a reason they rated a humor attempt as funny and, indeed, it is well known that exaggeration is commonly employed in humor (e.g., Kotthoff, 2006), including humor in Spanish (e.g., Shively, 2018). Simile and metaphor were also highlighted as reasons for a response being funny. Further, the limitation of the written DCT format was revealed in some comments made by raters; they pointed to written representations of prosodic, paralinguistic, and non-verbal information (e.g., sound lengthening, laughing, winking, eye rolling) as part of the reason that they thought a response was funny and highlighted that funniness may rest on the way an utterance was delivered, with the appropriate intonation and pauses. Eliciting oral responses to the DCT scenarios rather than written ones would be a solution to this problem while, at the same time, preserving the positive features of this type of instrument (e.g., ease of comparison). Finally, self-deprecating humor was well received by raters, perhaps because, as indicated in previous studies, it can make a person more approachable (e.g., Boxer & Cortés-Conde, 1997) and may demonstrate that a person has a good sense of humor and is able to laugh at one's own shortcomings and overcome them (e.g., Norrick, 1993).

In contrast, reasons raters gave for why humor was unsuccessful included not understanding the joke, not recognizing the humorous frame, and not having shared knowledge of a reference—all of which have previously been documented as reasons why humor can fail (Bell, 2015). Attempts at teasing the interlocutor in the DCT scenarios was a specific area with which many humor attempts were unsuccessful, either because the playful key of the comment did not come through to the raters and the utterance sounded like serious criticism or scolding or because the topic of the tease was perceived as too aggressive. Raters were not always in agreement about whether teasing was funny. Indeed, previous research suggests that teasing is a risky form of humor, since the hearer may be offended by being positioned as the butt of the joke (e.g., Boxer & Cortés-Conde, 1997). For this reason, teasing often occurs among people who know each other well, since in that case, the speaker tends to be better able to judge what teases will not be perceived as genuine insult and the hearer is less likely to misinterpret the playful intent and suffer a loss to face (e.g., Bell, 2009; Lampert & Ervin-Tripp, 2006). It is not possible to know whether, in real life, the participants in this study would choose to tease someone in the scenarios posed in the DCT, but the analysis here suggests that, if they did, many of their apparent attempts at playful teasing would run the risk of missing the mark.

Calques were another source of divergence among raters. For instance, the analysis indicated that translating literally the phrase "to have a green thumb" from English to Spanish was only funny to the L1-

English-speaking rater. However, it was not lack of familiarity with the phrase in English on the part of the L1-Spanish-speaking raters; they simply did not find the translation funny in Spanish. Calques such as this one are a reflection of the bilingual background of the participants and among bilinguals, creatively playing with linguistic rules is common (Vaid, 2006). It may be that certain types of bilingual humor are more likely to be appreciated by those who share their L1 or who have experience learning the same L2. Shively (2018) reported, for example, that one L2 speaker of Spanish who joked around with false cognates in English and Spanish only caused confusion, not amusement, for the L1 Spanish hearer.

5.1 Pedagogical implications

The findings reported here can inform instruction about humor and irony in Spanish. As part of a curriculum that includes these topics, instructors could employ the examples of successful and unsuccessful humor produced by L2 and heritage speakers, along with raters' responses, to assist in raising awareness among language learners about how different types of humor attempts may be perceived and why. Learners can further be advised that teasing is a more risky type of humor with more chance of failure if not targeted just right for the interlocutor and context and if the humorous frame is not made abundantly clear. It may also be insulting to the hearer if care is not taken to consider the topic of the teasing and whether the comment would be hurtful. Further, self-deprecating humor can be presented as a positive choice in that it has the potential to enhance a speaker's image by positioning him/herself as having a good sense of humor.

Through instruction, learners could also be made aware that sense of humor varies from one person to another, even within the same community or culture. Therefore, even if one's humor is not well received by one person, that does not necessarily mean that it will not be funny to another. While keeping this variation in mind, students can be made aware that there may be commonly shared perceptions by speakers of the L2/heritage language, such as the unanimous agreement among the L1-Spanish-speaking raters in this study that joking around by calling someone a *payaso* ('clown') was risky because it carries a strong possibility of offense. Previous research suggests that interventions to teach humor and irony in an L2 and heritage language can be successful in enhancing students' awareness (e.g., Shively et al., 2022; Kim & Lantolf, 2018; Petkova, 2013) and the findings reported here add to the resources that instructors have at their disposal to teach these two topics.

5.2 Limitations and future directions

As highlighted above, an important limitation of this study was the use of a written rather than oral DCT. Bell et al. (2021), who conducted their analysis contemporaneously to this study, came to the same conclusion with their written humor DCT. Hence, it would be beneficial for future research with elicited humor data to audio- or video-record oral productions of humor attempts, so that prosodic, paralinguistic, and non-verbal cues used by participants in their delivery of humor are captured and can be analyzed. The manner in which a humorous utterance is delivered may make the difference between a humor attempt being funny or not. To preserve the potential benefit of giving L2 and heritage speakers enough time to

craft their humor, while at the same time collecting spoken language data, participants in future studies could be given time to plan their response to a DCT scenario and then be asked to record it orally.

With regard to the raters who evaluated humor attempts in this study, they were a small group and their backgrounds were varied in terms of nationality, L1, and gender. Future research that involves a greater number of raters who share similar identities and backgrounds would be better positioned than the present study to be able to draw conclusions about the potential influence of demographic and linguistic variables on the rating of humor. Further, while the goal in this study was to examine ratings of humor attempts using data collected through a questionnaire, researchers in future studies could also recruit raters to evaluate naturally-occurring instances of humor or ask L2 and heritage speakers' interlocutors in naturalistic conversation to talk about, retrospectively, how they perceived naturally-occurring instances of humor in an interaction in which they were the intended recipient of the humor attempt.

Another limitation of the present study is the small number of participants, especially heritage speakers. In addition to including more L2 and heritage speakers in future research, the field would be well served by specifically focusing on humor produced by heritage speakers, an area that heretofore has not received much attention, yet that may tell a different story about humor practices in bilinguals.

6. Conclusion

Most previous research on both L2 and heritage speakers' humor practices has examined spontaneous instances of humor as they arise in naturalistic interactions. While such data offer a rich understanding of the multilayered ways in which bilinguals use humor to amuse and to achieve other interactional goals, this and a handful of other studies demonstrate that investigating bilingual humor through elicited data can also make contributions to our understanding in this area. As we gain increased understanding of bilingual humor, from both qualitative and quantitative approaches, we will be better positioned to not only appreciate the creativity and linguistic dexterity of bilingual humorists, but we will also be able to better develop language instruction focused on humor. Considering that humor is commonplace in everyday interactions and serves a variety of social functions that are crucial in building and maintaining relationships, as well as the fact that humor may make language learning more enjoyable and memorable (e.g., Bell, 2011), it behooves scholars and practitioners of language teaching and learning to take seriously—pun intended—how we can best support language learners to engage successfully with humor.

References

Adelswärd, V., & Öberg, B. (1998). The function of laughter and joking in negotiation activities. *Humor: International Journal of Humor Research*, *11*, 411-429.

Ahn, S. (2016). Exploring language awareness through students' engagement in language play. *Language Awareness*, *25*, 40-54.

Bell, N. D. (2006). Interactional adjustments in humorous intercultural communication. *Intercultural Pragmatics*, *3*, 1-28.

Bell, N. D. (2009). Responses to failed humor. *Journal of Pragmatics*, *41*, 1825-1836.

Bell, N. D. (2011). Comparing playful and nonplayful incidental attention to form. *Language Learning*, *62*, 236-265.

Bell, N. D. (2015). *We are not amused: Failed humor in interaction*. Berlin: De Gruyter.

Bell, N. D., & Attardo, S. (2010). Failed humor: Issue in non-native speakers' appreciation and understanding of humor. *Intercultural Pragmatics*, *7*, 423-447.

Bell, N. D., & Pomerantz, A. (2015). *Humor in the classroom: A guide for language teachers and educational researchers*. New York: Routledge.

Bell, N. D., & Pomerantz, A. (2019). Humor in L2 pragmatics research. In N. Taguchi (Ed.), *The Routledge handbook of second language acquisition and pragmatics* (pp. 63-77). New York: Routledge.

Bell, N. D., Shardakova, M., & Shively, R. L. (2021). The DCT as an assessment method for L2 humor production. In C. Félix-Brasdefer & R. L. Shively (Eds.), *New directions in second language pragmatics* (pp. 156-178). Berlin: De Gruyter.

Bell, N. D., Skalicky, S., & Salsbury, T. (2014). Multicompetence in L2 language play: A longitudinal case study. *Language Learning*, *64*, 72-102.

Belz, J. (2002). Second language play as a representation of the multicompetent self in foreign language study. *Journal of Language, Identity & Education*, *1*, 13-39.

Boxer, D., & Cortés-Conde, F. (1997). From bonding to biting: Conversational joking and identity display. *Journal of Pragmatics*, *27*, 275–294.

Carreira, M., & Hitchins Chik, C. (2018). Differentiated teaching: A primer for heritage and mixed classes. In K. Potowski (Ed.), *The Routledge handbook of Spanish as a heritage language* (pp. 359-374). New York: Routledge.

Chang, A. (2017). "Call me a little critical if you will": Counterstories of Latinas studying abroad in Guatemala. *Journal of Hispanic Higher Education*, *16*, 3-23.

Cheng, W. (2003). Humor in intercultural conversations. *Semiotica*, *146*, 287–306.

Chiaro, D. (2009). Cultural divide or unifying factor? Humorous talk in the interaction of bilingual, cross-cultural couples. In N. Norrick & D. Chiaro (Eds.), *Humor in interaction* (pp. 211-231). Amsterdam: John Benjamins.

Davies, C. E. (2003). How English-learners joke with native speakers: An interactional sociolinguistic perspective on humor as collaborative discourse across cultures. *Journal of Pragmatics*, *35*, 1361-1385.

Dynel, M. (2014). Isn't it ironic? Defining the scope of humorous irony. *Humor*, *27*, 619-639.

Dynel, M. (2017). Academics vs. American scriptwriters vs. academics: A battle over the etic and emic 'sarcasm' and 'irony' labels. *Language and Communication*, *55*, 69-87.

Dynel, M. (2018). *Irony, deception, and humor: Seeking the truth about overt and covert untruthfulness*. Mouton de Gruyter.

Forman, R. (2011). Humorous language play in a Thai EFL classroom. *Applied Linguistics*, *32*, 541-565.

Golato, A. (2003). Studying compliment responses: A comparison of DCTs and recordings of naturally occurring talk. *Applied Linguistics*, *24*, 90–121.

Kim, J., & Lantolf, J. P. (2018). Developing conceptual understanding of sarcasm in L2 English through explicit instruction. *Language Teaching Research*, *22*, 208-229.

Kotthoff, H. (2006). Pragmatics of performance and the analysis of conversational humor. *Humor*, *19*, 271-304.

Kuipers, G. (2015). *Good humor, bad taste: A sociology of the joke*. Berlin: Mouton de Gruyter.

Lampert, M. D., & Ervin-Tripp, S. M. (2006). Risky laughter: Teasing and self-directed joking among male and female friends. *Journal of Pragmatics*, *38*, 51-72.

Leung, G. (2014). Hoisan-wa in jest: Humor, laughter, and the construction of counter-hegemonic affect in contemporary Chinese American language maintenance. *Humor*, *27*, 203-225.

Martínez, R. A., & Morales, P. Z. (2014). ¿*Puras groserías*?: Rethinking the role of profanity and graphic humor in Latin@ students' bilingual word play. *Anthropology & Education Quarterly*, *45*, 337-354.

Norrick, N. (1993). *Conversational joking: Humor in everyday talk*. Bloomington, IN: Indiana University Press.

Ogiermann, E. (2018). Discourse completion tasks. In A. H. Jucker, K. P. Schneider & W. Bublitz (Eds.), *Methods in pragmatics* (pp. 229–255). Berlin: De Gruyter.

Padilla García, X. (2008). Claves para abordar la ironía en programas de ELE: De la competencia comunicativa a la competencia intercultural. *VERBA*, *35*, 275-303.

Petkova, M. (2013). *Effects and perceptions of a humor competence curriculum in an intensive English program in Southern California*. Unpublished doctoral dissertation, Alliant International University, San Diego, CA.

Potowski, K. (2018). Spanish as a heritage/minority language: A multifaceted look at ten nations. In K. Potowski (Ed.), *The Routledge handbook of Spanish as a heritage language* (pp. 1-14). New York: Routledge.

Ruch, W. (2008). Psychology of humor. In V. Raskin (Ed.), *The primer of humor research* (pp. 17-100). Berlin: de Gruyter.

Shardakova, M. (2013). "I joke you don't": Second language humor and intercultural identity construction. In C. Kinginger (Ed.), *Social and cultural aspects of language learning in study abroad* (pp. 207-238). Amsterdam: John Benjamins.

Shively, R. L. (2018). *Learning and using conversational humor in a second language during study abroad*. Berlin: de Gruyter.

Shively, R. L., Acevedo, J., Cano, R., & Etxeberria-Ortego, I. (2022). Teaching humorous irony to L2 and heritage speakers of Spanish. *Language Teaching Research*, *26*, 279-302.

Taguchi, N. (2011). Rater variation in the assessment of speech acts. *Pragmatics*, *21*, 453-471.

Vaid, J. (2006). Joking across languages: Perspectives on humor, emotion, and bilingualism. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression, and representation* (pp. 152-182). Clevedon, UK: Multilingual Matters.

Vandergriff, I., & Fuchs, C. (2012). Humor support in synchronous computer-mediated classroom discussions. *Humor*, *25*, 437-458.

Zentella, A. C. (2003). "José can you see?": Latin@ responses to racist discourse. In D. Sommer (Ed.), *Bilingual games: Some literary investigations* (pp. 51-69). New York: Macmillan.

Bionote

Rachel Shively (PhD, University of Minnesota) is a Professor of Spanish and Applied Linguistics in the Department of Languages, Literatures, and Cultures at Illinois State University. Shively's research focuses on second language pragmatics, second language humor, and study abroad. She is currently editor of the journal Study Abroad Research in Second Language Acquisition and International Education.