Incremental Interpretation in the First and Second Language

Laurel Brehm, Carrie N. Jackson, and Karen L. Miller

1. Introduction
1.1. Background

In conversation, listeners are sometimes faced with trying to interpret erroneous utterances such as '*Nobody can't say we weren't fair*' or dialectal variants such as *'There's four of them'*. Anecdotally, it seems that we can do this easily in our first language (L1) and, with some attainment of proficiency, can also do so in our second-language (L2). However, less is known about how these skills develop: How do we learn to process utterances non-literally? In the current study, we examine this question by comparing how L1 English speakers and proficient L1 Spanish-L2 English speakers process and interpret anomalous utterances. In so doing, this study addresses open questions regarding how L1 and L2 speakers deal with conflicting linguistic cues in comprehension.

Two frameworks provide a set of proposed mechanisms by which individuals can understand speech non-literally. These are the *noisy channel* (e.g. Gibson, Bergen, & Piantadosi, 2013; Gibson et al, 2017; Levy, 2008) and *good-enough processing* frameworks (e.g. Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Ferreira, 2003). The notion behind both frameworks is that since the goal of speech is typically to communicate a message from a speaker to a listener, listeners often interpret anomalous or unpredictable utterances non-literally, inferring what was likely to be intended rather than strictly sticking to what was actually said. Under the noisy channel framework, non-literal inferences are drawn by mentally comparing observed utterances to others that are more likely and differ minimally in terms of insertions/deletions of sounds or letters; under the good-enough processing framework, non-literal inferences are drawn from heuristics or biases about likely words, structures, or events. In either case, by relying on expectations about the likelihood of words and structures in language, listeners are provided resilience in the face of produced errors, noise in the system, and any unknown dialectal variants, making communication robust across speakers and situations.

An implication that follows from both non-literal processing frameworks is that L2 speakers may differ from L1 speakers in processing: both frameworks

* Laurel Brehm, Max Plank Institute for Psycholinguistics, laurel.brehm@mpi.nl, Carrie N. Jackson and Karen L. Miller, Pennsylvania State University. This work was supported by a grant from the National Science Foundation (NSF OISE 0968369; PI: J. F. Kroll; co-PIs: P. E. Dussias and J. G. van Hell). Thanks to Alison Kelly, Jennifer Rhee, and Katherine Muschler for assistance with data collection.

 Proceedings of the 43rd Boston University Conference on Language Development, ed. Megan M. Brown and Brady Dailey, 109-122. Somerville, MA: Cascadilla Press.

rely on expectations about others' speech. Furthermore, other L2 processing frameworks suggest that L1 and L2 speakers may differ fundamentally in how utterances are processed. This means that an examination of non-literal processing in the L2 informs how anomaly processing develops and provides insight about how L1 and L2 processing might or might not fundamentally differ.

1.2. Non-literal Processing in the L1 and L2

Three related but distinct frameworks provide hypotheses for how non-literal processing might operate differently in the L1 and L2.

One possibility, derived from the noisy-channel and good-enough processing frameworks, is that L1 and proficient L2 speakers might process anomalies differently due to their different sets of expectations about utterance likelihood, leading to differences in online and offline processing. Under this hypothesis, more non-literal inferences are expected for L2 than L1 speakers because L2 speakers have a less-precise noise model due to their reduced experience with the language (e.g. Futrell & Gibson, 2017). A potential consequence of this during online processing is that alternate parses of utterances may be more compelling, causing L2 speakers to receive more interference from locally-plausible cues that are not compatible with the literal sentence structure.

Another possibility is that L2 processing might be inherently shallower and more non-literal than L1 processing, which would also lead to differences in online and offline processing. This hypothesis has been discussed in the literature as the *Shallow Structure Hypothesis* (SSH, see Clahsen & Felser, 2006). Shallow processing leads to a qualitative difference in the outcomes and trajectories of processing for L2 and L1 individuals: for example, when processing semantically implausible sentences, L2 learners overuse semantic heuristics, vs literal syntax (e.g. Lim & Christianson, 2013). The prediction here is that L2 speakers will be more likely to process utterances non-literally, regardless of whether they are anomalous. In addition, a strong version of the SSH suggests that individuals may be less able to revise their interpretation of an utterance's structure, leading to less attention to locally-plausible cues.

A third possibility is that L2 processing might be inherently resource-intensive, but is not qualitatively different than L1 processing. A recent framework (Cunnings, 2017) has attributed a number of L2 processing phenomena to working memory difficulty. The proposal is that given that the words, structures, and morphology of the L2 are, by definition, less well-rehearsed than those in the L1, they may be harder to retrieve from memory. The resulting memory burden of processing in the L2 then may lead to more reliance on recent cues and an increase in non-literal inferences. As the fundamental underlying processing mechanisms remain the same for L1 and L2 speakers, the implication is that L1 and L2 speakers should show highly similar processing strategies, with similar amounts of attention to referents but that L2 speakers may be slower to switch attention away from a referent once attention has been allocated to it due to retrieval difficulty and increased overall interference.

The current work tests these hypotheses by examining how subject-verb agreement anomalies are processed and understood by proficient L2 speakers of English. We ask whether L2 speakers differ from L1 speakers in the rate of non-literal inferences about agreement or in incremental agreement processing.

1.3. Subject-Verb Agreement

We examined subject-verb agreement as it shows a large amount of variability in production and comprehension: errors and anomalies in agreement are frequently observed in real-world speech, yet such anomalies have also been shown to be relatively easy to process. The pattern in production is that fragments with a singular head and plural local noun like "*The key to the cabinets*" are often completed with a plural verb (1a), instead of the canonical completion (1c). This pattern obtains in English (e.g. Bock & Miller, 1991) and in Spanish (e.g. Vigliocco, Butterworth, & Garrett, 1996). In processing, sentences containing agreement-anomalies are read with no apparent slow-down for items like 1a vs 1c (Wagers, Lau, & Phillips, 2009). Similar patterns are elicited from proficient L2 speakers of English in production (Nicol & Greth, 2003; Jackson, Mormer & Brehm, 2018) and processing (Lim & Christianson, 2015). This suggests that L1 and proficient L2 speakers of English process subject-verb agreement similarly: plural cues increase the production of plural agreement and reduce the cost of processing ungrammatical plural verbs.

1a. The key to the cabinets literally *were on the table
1b. The key to the cabinet literally *were on the table
1c. The key to the cabinets literally was on the table
1d. The key to the cabinet literally was on the table

For L1 readers, the ultimate interpretation of agreement anomalies has also been explicitly tested. The finding is that sentences containing plural cues on local (non-subject) nouns or on verbs often lead to a non-literal interpretation of the head number, such that 1a-c all lead to the non-literal inference that the head was *keys* and not the literal *key* (e.g. Patson & Husband, 2016; Brehm, Jackson, & Miller, 2018). The question we ask in the current work is how non-literal interpretations of subject-verb agreement arise for L2 speakers, and how this informs the time-course and mechanisms of non-literal processing. We compare L1 Spanish-L2 English speakers to an L1 English group tested in an earlier experiment; the results of this L1 group appear in the following section.

1.4. Incremental L1 Agreement Processing

In earlier work (Brehm, Jackson, & Miller, in prep), we presented individuals with four-image arrays consisting of single and plural versions of the head and local nouns. We tracked participants' eyes while they viewed these images and listened to sentences like 1a-d. After listening to the sentence, participants selected which image matched the subject of the sentence. This provided us with an offline measure of interpretation and an online measure of processing.

We showed that, as in Patson and Husband (2016) and Brehm et al (2018), non-literal interpretations of head number (*key* → *keys*) were more common if the sentence contained plural local nouns and/or anomalous plural verbs (e.g. 1a-c). Non-literal interpretations also decreased through the course of the experiment, especially for sentences containing plural verbs. This shows the quick adaptation of L1 speakers to this type of anomaly.

For sentences interpreted literally, we also examined the fixations made to a non-literal version of the head differing in number ('keys'). This indexes whether participants were considering an alternate parse of the sentence at various points during processing, even when the participant ended up interpreting the sentence literally. Mirroring the offline interpretation data, we showed that both types of number cues—plural nouns and anomalous plural verbs—led participants to fixate a plural version of the head noun ('keys'). In the time window starting 200 msec from verb onset, there was a significant main effect of local noun type, such that sentences with plural local nouns (1a, 1c) elicited more fixations to the non-literal head. In the time window starting 700 msec from verb onset, there was a significant main effect of verb type, such that sentences with plural verbs (1a, 1b) had more fixations to the non-literal head. Fixations to the non-literal head also decreased throughout the course of the experiment in both regions and the region beginning 200 msec from the onset of the noun, providing evidence for L1 participants' adaptation to anomalies.

These data show that when processing agreement, L1 English speakers consider a version of the head noun differing in number if any noun or verb inflections mismatch the head. This leads to more non-literal interpretations and to increased attention to a non-literal version of the head after hearing nouns and verbs with plural inflections, even when the sentence was interpreted literally.

1.5. Current Study

The current study assesses whether these observed patterns replicate among proficient L1 Spanish-L2 English speakers. All accounts suggest that L2 speakers are likely to have a higher rate of non-literal inferences than L1 speakers. This might be because of a more variable noise model (under the noisy-channel theory or good-enough processing), because of a reliance on shallow cues (under the SSH), or because of memory difficulty.

What is diagnostic, however, is the comparison of online processing across L1 and L2 speakers who have successfully interpreted an utterance literally. One version of the noisy-channel theory predicts that L2 speakers might be more distracted by local cues, leading to more attention to locally-plausible cues. In contrast, if L2 speakers form a shallow parse of the input, one prediction would be that they are unable to revise their initial interpretation. This might mean that on trials where the utterance was interpreted literally, L2 speakers will show less attention to the plural version of the head than L1 speakers. Finally, if L2 speakers deploy representations more slowly than L1 speakers, as suggested by a memory deficit framework (e.g., Cunnings, 2017), then the L2 group may show attention to alternate referents at later time windows than L1 speakers, even when the sentence is ultimately interpreted in the same way. In addition, memory deficits may lead L2 speakers to have more trouble inhibiting competing cues, which would lead to more attention to the non-literal head.

2. Method

2.1. Participants

Data were collected from 33 individuals recruited from the Pennsylvania State University community via message boards, international student groups, and word-of-mouth. Participants were paid $15 in exchange for participation. Four individuals were excluded for patterning as early bilinguals, and one was excluded due to eye-tracking calibration difficulty. The 28 remaining participants were native speakers of Latin American Spanish, ranged in age from 19 to 38 (M=23.8) and had normal or corrected-to-normal vision and hearing. Demographic data can be found in Table 1.

Table 1. *Demographics of participant sample. SD in parentheses.*

Age	*M*=23.8 (4.6) Range 19-38
Level of Education	At least some college: 17 At least some graduate: 11
Reported native language	Spanish (100%)
% English use during typical day	*M*=39.1% (27.5%) Range 4-98%
English Proficiency (MELICET)	*M*= 81% (12%) Range 58-98%
English Age of Acquisition	*M*= 7.4 (3.8) Range 4-15
Years of English	*M*= 15.7 (5.1) Range 7-27

2.2. Materials and Design

The 48 critical items were all taken from previous work (Brehm et al, in prep). Items had a singular head and the number of the local noun (singular/plural) and verb (singular/plural) were crossed to make four versions of each item, as shown in sentences 1a-d above. Critical items were paired with 60 filler items, 32 of which contained semantically-odd ‘Without’ blends (as described in Brehm et al, 2018; Frazier & Clifton, 2015). The remaining 28 filler items used datives, began with numerals, and contained embedded relative clauses.

Critical and filler items were recorded by a female speaker of American English from the Washington DC area. Critical item recordings were cross-spliced before the anomalous verb in order to match versions of the item for acoustic content up until the point of the anomaly, such that items like 1a and 1c. as well as 1b and 1d, began with identical recordings.

All items were paired with a grid of four colored line drawings compiled from stimulus databases (Duñabeitia et al, 2017; Rossion & Pourtois, 2004), public domain sources (Wikimedia Commons, Pixabay, and Flickr), created using Pixton.com, or drawn by the first author. The image grid for critical trials always contained one image of a single token of the head noun, one image of two tokens of the head noun, one image of a single token of the local noun, and one image of two tokens of the local noun (see Figure 1). Filler items used images of cartoon people, indoor and outdoor locations, animals, and objects.

Critical items were assigned to 16 lists using a Latin square design such that the four versions of the items were equally represented in a list and so that the target image was equally likely to appear in each screen location. Within a list, items were presented in a fixed pseudo-randomized order with an equal number of items of each condition assigned to each of four blocks, and with no more than two critical items (and no two critical items in the same condition) adjacent.

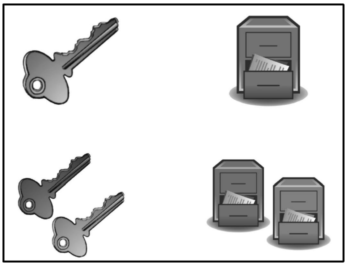

Figure 1. Sample image array for sentence 1.

2.3. Apparatus and Procedure

Participants were instructed to listen to recorded sentences and to select the image from the array on the screen that best matched the subject of the sentence. The concept of 'Subject' was defined with the statement: "The subject of the sentence is the do-er of an action or the thing that is being described".

Trials began with a fixation target in the middle of the screen; participants clicked on this to start the trial. Next, the image array appeared for 1500 msec, which served as preview time. Then, a recorded sentence was played over the computer speakers, which lasted between 2213 and 3495 ms. Finally, participants used the mouse to select the image that matched the sentence subject; they were given an unlimited amount of time to do so. An orange box appeared around the selected image and remained for 2000 msec.

The eye-tracking experiment was run using Experiment Builder on a Dell PC with a 21 x 11.5 inch monitor and an EyeLink 1000 Plus Desktop in remote mode (sample rate 500 Hz) that was 20 inches in front of the computer monitor. Participants were seated 20 inches in front of the camera. Images appeared in ports, defining the experimental interest areas. These were 6.75 inches x 5 inches, spaced 2.75 inches apart horizontally and 1.5 inches apart vertically.

After running the eye-tracking experiment, participants completed a language background questionnaire and the MELICET English proficiency exam (used with permission from Blattner, 2007). They were then debriefed and given the opportunity to ask questions. The session lasted about 75 minutes.

2.4. Analysis

All analyses were conducted using R (version 3.3.3, R Core Team, 2017) using the package lme4 (version 1.1-13, Bates, Maechler, Bolker, & Walker, 2015). Analysis of subject selections used a logistic regression on the odds of selecting a non-literal target object (an object other than 'key'). In this analysis, predictors were Local Noun (singular, plural, contrasts of .5, -.5), Verb (singular, plural, contrasts of .5, -.5), and Trial Number (scaled and centered). In this model, no interaction was entered between Local Noun and Verb due to the few non-literal responses in Singular Local Noun, Singular Verb trials.

The eye-tracking analysis used a linear regression on the proportion of fixations per person per trial to the non-literal head object ('*keys*'). Separate analyses were done on four 500 msec time bins. These were offset by 200 msec from the onset of critical words (head noun, local noun, and verb), with an additional window beginning 700 msec after the onset of the verb. We analyzed only trials in which the literal subject 'key' was selected, reflecting the dominant response type. Predictors were Local Noun (singular/plural, contrasts of .5, -.5 centered separately per region[1]), Verb (singular/plural, contrasts of .5, -.5 centered separately per region) and Trial Number (linear; scaled and centered).

For all analyses, random intercepts were included for Participants and Items and as many random slopes as justified by the data were included (e.g. as suggested by Barr, Levy, Scheepers, & Tily, 2013). Random slopes were removed due to non-convergence (beginning with the highest-order interactions) or due to correlations above 0.9 between random terms.

3. Results

3.1. Sentence Interpretations

As shown in Figure 2, the dominant interpretation of all items was the literal one (the subject was *key)*, regardless of whether the item contained a plural local noun or an anomalous plural verb. Non-literal interpretations (the subject was

---

[1] Centering a two-level contrast is done by subtracting the mean value from each contrast. This gives cells with unequal observations equivalent weight statistically.

keys) increased for sentences containing plural local nouns and anomalous plural verbs, with the most non-literal plural interpretations of the head noun occurring in the local plural noun, plural verb condition. These patterns were supported by logistic mixed-effect regression models (see Table 2) with significant main effects of local noun and verb and an interaction between the two.

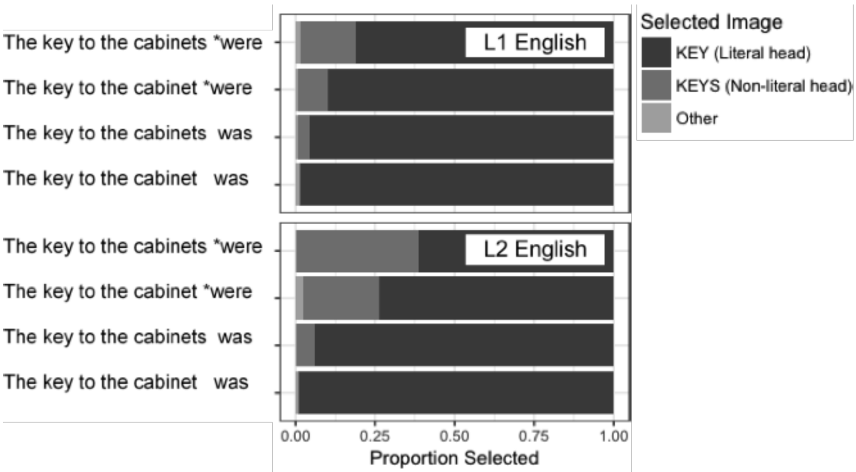

Figure 2. Non-literal subject-verb agreement interpretations for L2 English speakers, split by interpretation and sentence type. L1 data from Brehm et al (in prep) appears in top panel for comparison.

Table 2. *Logistic mixed-effect regression for odds of non-literal interpretations.*

Fixed Effects	$\hat{\beta}$	SE	z value	p value
Intercept	-4.46	0.66	-6.70	<0.001
Local Noun	-3.67	0.91	-4.02	**<0.001**
Verb	-5.50	0.84	-6.54	**<0.001**
Trial Number	-0.12	0.18	-0.69	0.49
Noun x Verb	-3.47	1.28	-2.71	**0.01**
Noun x Trial	0.18	0.25	0.72	0.47
Verb x Trial	-0.54	0.33	-1.62	0.11
Random Effects		*Term*	*SD*	
	Item	Intercept	1.09	
		Noun	0.82	
		Verb	1.49	
	Participant	Intercept	2.18	
		Noun	1.92	

3.1.1. Comparison with L1 Group

We performed omnibus analyses that compared the L2 group with a group of L1 English speakers performing an identical task (Brehm et al, in prep) in order to see whether L2 speakers differed quantitatively from L1 speakers. Experiment was added as a main effect and an interaction with all other predictors (contrast coded: L1= -0.5, L2=0.5).

We observed a main effect of Experiment, such that the L2 group had more misinterpretations than the L1 group. We also observed a reliable interaction between Experiment and Trial Number, such that the L1 group tended to have fewer non-literal interpretations toward the end of the experiment, while the L2 group had a constant rate of non-literal interpretations across the experiment. No other effects including Experiment were reliable.

3.2. Fixation Patterns on Literally-Interpreted Items

For all trials interpreted literally, we examined the likelihood of fixating on a number competitor for the head noun ('keys') over the time-course of the trial. This is shown in Figure 3. We performed analyses in four time windows related to critical words in the sentence. In the first analysis window, reflecting processing from 200 to 700 msec after the onset of the head noun there were no significant effects (see Table 3). In the second analysis window, reflecting processing from 200 to 700 msec after the onset of the local noun, we observed an interaction between Local Noun and Verb. This was reflected in a cross-over interaction, such that trials with a local plural noun and a plural verb elicited more looks to the non-literal referent than local plural, singular verb trials, while trials with a local singular noun and a plural verb elicited more looks to the non-literal referent than local singular, singular verb trials (see Figure 3). This pattern is likely to be due to chance, due to matching sound files across items. No other effects were reliable (see Table 3).

In the regions following the anomalous verb, there were larger observed differences in looks to the non-literal referent. In the third analysis window, reflecting processing from 200 to 700 msec after the onset of the verb, we observed a main effect of Local Noun, such that trials with a plural local noun elicited more looks to the non-literal referent (*keys*) than trials with a singular local noun (see Figure 3). No other effects were reliable (see Table 3). In the fourth region, representing processing from 700 to 1200 msec after the onset of the verb, trials containing a plural verb elicited more looks to the non-literal referent (*keys,* see Figure 3). No other effects were reliable (see Table 3).

3.2.1. Comparison with L1 Group

At each analysis window, we again ran omnibus analyses comparing the L2 group with a group of L1 English speakers performing the same task (Brehm et al, in prep). We again added Experiment as a main effect and an interaction with all other predictors (contrast coded: L1= -0.5, L2=0.5 and then centered for all regions). Analyses used the same analysis windows described above.

There were no reliable differences involving Experiment in the first two analysis windows (reflecting the head noun and local noun). However, differences appeared in the later analysis windows, reflecting processing in response to the anomalous verb. In the third window, beginning 200 msec from the onset of the verb, there were interactions between Experiment and Trial Number, and between Experiment, Local Noun Type, and Trial Number. The pattern was that the L2 group tended to look more at the non-literal referent (*keys*) when the local noun was plural, especially in the later portions of the experiment. In the fourth window, beginning 700 msec from the onset of the verb, there was an interaction between Experiment and Local Noun type. The pattern was that in the L2 group, there was a lingering effect of local noun type, such that a local plural noun still led to more looks to the non-literal referent (*keys*) compared to the L1 group. There was also an interaction between Experiment and Trial number, such that the L1 group looked less at the non-literal referent (*keys*) in trials late in the experiment, while the L2 group was relatively constant in their fixations the non-literal referent across trials.

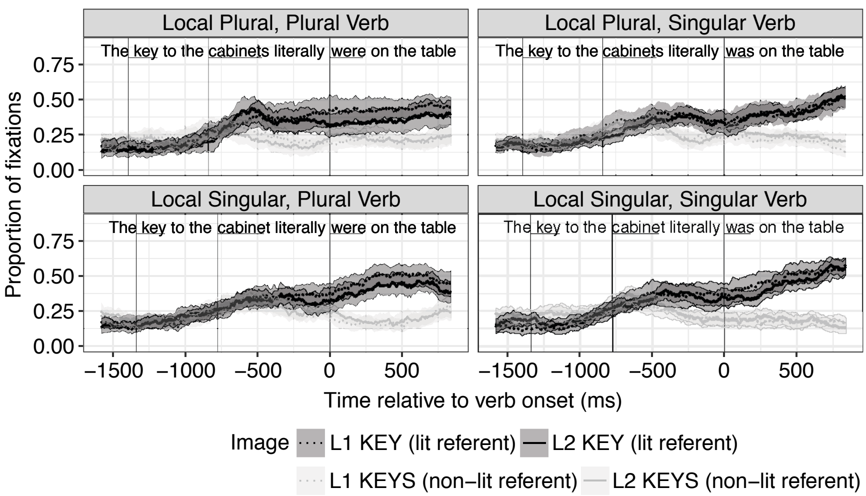

Figure 3. Fixations to literal (black, 'key') and non-literal (grey, 'keys') referent over time (zeroed to verb onset) for sentences interpreted literally. Panels reflect stimulus versions. Solid lines reflect L2 group: L1 group (from Brehm et al, in prep) is plotted in dashed lines for comparison. Vertical lines reflect average onsets of critical words within the sentence. Confidence bands are 95% CIs from a non-parametric bootstrap (1000 iterations) sampled over participants with replacement at 10 msec intervals.

Table 3. Mixed-effect model results for eye-tracking analyses by region. All p values were calculated via model comparison.

	Head noun region *Onset of head noun + 200 to Onset of head noun +700*				**Local noun region** *Onset of local noun + 200 to Onset of local noun +700*				**Verb region** *Onset of verb + 200 to Onset of verb +700*				**Post-Verb region** *Onset of verb + 700 to Onset of verb +1200*			
Fixed Effects	$\hat{\beta}$	*SE*	*t val.*	*p val.*	$\hat{\beta}$	*SE*	*t val.*	*p val.*	$\hat{\beta}$	*SE*	*t val.*	*p val.*	$\hat{\beta}$	*SE*	*t val.*	*p val.*
Intercept	0.51	0.02	29.25	< 0.001	0.25	0.01	19.38	< 0.001	0.21	0.01	14.65	< 0.001	0.19	0.01	14.19	< 0.001
Local Noun	0.00	0.03	0.18	0.86	0.03	0.02	1.35	0.18	-0.05	0.02	-2.13	**0.04**	-0.04	0.02	-1.79	0.08
Verb	-0.01	0.03	-0.55	0.58	0.02	0.02	0.81	0.42	0.02	0.02	0.90	0.37	-0.09	0.02	-4.84	**< 0.001**
Trial Number	0.00	0.01	0.18	0.86	0.01	0.01	0.50	0.62	-0.01	0.01	-0.92	0.36	0.00	0.01	-0.37	0.71
Noun x Verb	0.01	0.05	-0.22	0.94	-0.09	0.04	-2.21	**0.03**	0.01	0.04	0.15	0.88	-0.01	0.04	-0.36	0.72
Noun x Trial	0.03	0.03	1.26	0.66	0.00	0.02	-0.05	0.96	-0.03	0.02	-1.67	0.10	-0.02	0.02	-1.10	0.27
Verb x Trial	0.04	0.03	1.58	0.40	0.00	0.02	0.10	0.92	0.00	0.02	-0.08	0.94	0.01	0.02	0.47	0.64
Noun x Verb x Trial	-0.04	0.05	-0.65	0.67	0.06	0.04	1.40	0.16	-0.03	0.04	-0.74	0.46	0.02	0.04	0.49	0.63

Random Effects		Term	SD		Term	SD		Term	SD		Term	SD
	Item	Intercept	0.05	Item	Intercept	0.03	Item	Intercept	0.06	Item	Intercept	0.04
	Partic.	Intercept	0.05	Partic.	Intercept	0.04		Noun	0.09	Partic.	Intercept	0.04
	Resid.		0.29	Residual		0.33	Partic.	Intercept	0.04		Noun	0.04
							Residual	0.32		Residual	0.31	

4. Discussion

Subject-verb agreement is often mis-produced; consequently L1 English speakers are skilled at interpreting agreement anomalies non-literally. In the current work, we show that proficient L1 Spanish-L2 English speakers can interpret subject-verb agreement non-literally just like L1 English speakers do, 'listening through' observed anomalies to derive an interpretation that was not literally spoken. Like L1 English speakers, proficient L1 Spanish-L2 English speakers are more likely to interpret utterances that contain plural non-subject nouns or plural verbs as if they had contained a plural head noun. Like L1 English speakers, proficient L1 Spanish-L2 English speakers allocate attention to non-literal plural versions of the head noun upon encountering plural cues in the sentence conveyed upon nouns and verbs.

Despite the broad similarities, there were differences in the rate of non-literal interpretations between the two groups. This is predicted from all the frameworks discussed in the introduction. The L2 group had a 10% higher rate of non-literal inferences than the L1 group. This indicates that even for highly-proficient L2 speakers, utterances may not necessarily mean the same thing they do to an L1 speaker. This is consistent with all hypotheses raised in the introduction, and could result from the SSH (Clahsen & Felser, 2006), from the conjecture that L2 speakers may have a 'noisier' model of the speech around them due to reduced experience with their L2 (Futrell & Gibson, 2017), and/or from a memory deficit account of L2 processing (Cunnings, 2017).

Despite the robust differences in the rate of non-literal interpretations across groups, it was the case that for sentences interpreted in the same way (literally), differences between L1 and L2 speakers in incremental processing were much more minimal. We found that both the L1 and L2 groups incrementally and probabilistically considered a non-literal interpretation of the head's number after conflicting number cues appeared, with similar patterns of incremental processing for both groups when the same interpretation was obtained. This suggests that all listeners—L1 and L2 alike—consider a revision of a sentence's meaning probabilistically when there is evidence to do so. The implication is that L1 and L2 speakers fundamentally process utterances incrementally using similar mechanisms.

The general similarities between groups underscores the fact that L2 processing is quantitatively, not qualitatively different from L1 processing, providing evidence for the noisy-channel/good-enough processing framework and the memory-deficit account, as both of these accounts posit that similar mechanisms underlie L1 and L2 processing. However, these results do not support the SSH (Clahsen & Felser, 2006): The fact that similar processing patterns obtained for both groups when the same meaning was inferred further shows that proficient L2 comprehension can be equally specified, and equally deep relative to L1 comprehension.

Differences in incremental processing between the L1 and L2 groups disclose two general patterns: relative difficulty for L2 speakers in adapting during the experiment and a lingering influence of competing noun number for L2 speakers. This supports aspects of the noisy-channel framework and the memory deficit account of L2 processing. First, as shown in online and offline measures, L1 speakers tended to adapt more within the context of the experiment than the L2 speakers. Consistent with the noisy-channel framework, this suggests that L1 speakers may have a more refined 'noise' model and can therefore more easily adapt to the language statistics in the local environment. We also observed that the L2 speakers were more influenced by the local noun's number, and this influence persisted longer during processing. Consistent with the memory deficit account, this suggests that the local noun's number either remained active in L2 speaker's mental representations for longer or that it was more distracting, consistent with an effortful deployment of items from memory in the L2 (e.g. Cunnings, 2017).

5. Conclusion

Data from online measures of processing and offline measures of interpretation show that highly proficient L2 speakers of English are skilled in processing anomalies non-literally, patterning much like L1 speakers of English. L2 speakers are more likely to infer a non-literal interpretation, consistent with the noisy-channel framework, but when the same utterance meaning is construed, L1 and L2 speakers process utterances incrementally in nearly the same way. This suggests that differences in L1 and L2 comprehension do not rely on fundamentally different mechanisms, but that L2 speakers rely on the same mechanisms deployed in a noisier system that requires more effort.

References

Barr, Dale J., Levy, Roger, Scheepers, Christoph, & Tily, Harry J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278.

Bates, Douglas, Maechler, Martin, Bolker, Benjamin, & Walker, Steve (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014.

Blattner, Geraldine (2007). *Processing verbal arguments in a first and second language: The role of immersion experience* (Unpublished doctoral dissertation). Pennsylvania State University, University Park, PA.

Bock, Kathryn, & Miller, Carol A. (1991). Broken agreement. *Cognitive Psychology*, *23*(1), 45-93.

Brehm, Laurel, Jackson, Carrie N., & Miller, Karen L. (2018). Speaker-specific processing of anomalous utterances. *Quarterly Journal of Experimental Psychology*

Brehm, Laurel, Jackson, Carrie N., & Miller, Karen L. (in prep). Probabilistic online processing of sentence anomalies.

Christianson, Kiel, Hollingworth, Andrew, Halliwell, John F., & Ferreira, Fernanda (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, *42*(4), 368-407.

Clahsen, Harald, & Felser, Claudia (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(1), 3-42.

Cunnings, Ian (2017). Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*, *20*(4), 659-678.

Duñabeitia, Jon A., Crepaldi, Davide, Meyer, Antje S., New, Boris, Pliatsikas, Christos, Smolka, Eva, & Brysbaert, Marc (2017). MultiPic: A standardized set of 750 drawings with norms for six European languages. *The Quarterly Journal of Experimental Psychology*, 1-24.

Ferreira, Fernanda (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*(2), 164-203.

Frazier, Lyn, & Clifton Jr, Charles (2015). Without his shirt off he saved the child from almost drowning: interpreting an uncertain input. *Language, Cognition and Neuroscience*, *30*(6), 635-647.

Futrell, Richard, & Gibson, Edward (2017). L2 processing as noisy channel language comprehension. *Bilingualism: Language and Cognition*, *20*(4), 683-684.

Gibson, Edward, Tan, Caitlin, Futrell, Richard, Mahowald, Kyle, Konieczny, Lars, Hemforth, Barbara, & Fedorenko, Evelina (2017). Don't underestimate the benefits of being misunderstood. *Psychological Science*, *28*(6), 703-712.

Gibson, Edward, Bergen, Leon, & Piantadosi, Steven T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences*, 201216438.

Jackson, Carrie N., Mormer, Elizabeth, & Brehm, Laurel (2018). The production of subject-verb agreement among Swedish and Chinese second language speakers of English. *Studies in Second Language Acquisition*, 1-15.

Levy, Roger (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126-1177.

Lim, Jung Hyun & Christianson, Kiel (2013). Second language sentence processing in reading for comprehension and translation. *Bilingualism: Language and Cognition*, *16*(3), 518-537.

Lim, Jung Hyun & Christianson, Kiel (2015). Second language sensitivity to agreement errors: Evidence from eye movements during comprehension and translation. *Applied Psycholinguistics*, *36*(6), 1283-1315.

Nicol, Janet, & Greth, Delia (2003). Production of subject-verb agreement in Spanish as a second language. *Experimental Psychology*, *50*(3), 196.

Patson, Nikole D., & Husband, E. Matthew (2016). Misinterpretations in agreement and agreement attraction. *The Quarterly Journal of Experimental Psychology*. 69(5), 950-971

R Core Team (2017). R: a Language and Environment for Statistical Computing. Version 3.3. 3. R Foundation for Statistical Computing, Vienna, Austria.

Rossion, Bruno, & Pourtois, Giles (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*(2), 217-236.

Vigliocco, Gabriella, Butterworth, Brian, & Garrett, Merrill F. (1996). Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, *61*(3), 261-298.

Wagers, Matthew W., Lau, Ellen F., & Phillips, Colin (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*(2), 206-237.

Proceedings of the 43rd annual Boston University Conference on Language Development

edited by Megan M. Brown
and Brady Dailey

Cascadilla Press Somerville, MA 2019

Copyright information

ISSN 1080-692X
ISBN 978-1-57473-096-8 (2 volume set, paperback)

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com