Research Article

Psychological Science
1–10

DOI: 10.1177/0956797617690277
www.psychologicalscience.org/PS

Don't Underestimate the Benefits of Being Misunderstood

Edward Gibson[1], Caitlin Tan[1], Richard Futrell[1], Kyle Mahowald[1], Lars Konieczny[2], Barbara Hemforth[3], and Evelina Fedorenko[1,4,5]

[1]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology; [2]Center for Cognitive Science, University of Freiburg; [3]Centre National de la Recherche Scientifique (CNRS), Université Paris Diderot; [4]Harvard Medical School, Boston, Massachusetts; and [5]Massachusetts General Hospital, Boston, Massachusetts

Abstract

Being a nonnative speaker of a language poses challenges. Individuals often feel embarrassed by the errors they make when talking in their second language. However, here we report an advantage of being a nonnative speaker: Native speakers give foreign-accented speakers the benefit of the doubt when interpreting their utterances; as a result, apparently implausible utterances are more likely to be interpreted in a plausible way when delivered in a foreign than in a native accent. Across three replicated experiments, we demonstrated that native English speakers are more likely to interpret implausible utterances, such as "the mother gave the candle the daughter," as similar plausible utterances ("the mother gave the candle to the daughter") when the speaker has a foreign accent. This result follows from the general model of language interpretation in a noisy channel, under the hypothesis that listeners assume a higher error rate in foreign-accented than in nonaccented speech.

Keywords

comprehension, language, psycholinguistics, open data, open materials

Received 7/27/16; Revision accepted 12/29/16

Being a nonnative speaker of a language poses challenges. Individuals often feel embarrassed by their accents and the errors they make when speaking in their second language (Gluszek & Dovidio, 2010). Indeed, compared with native speakers, individuals with foreign accents are perceived to be less credible (Bourdieu & Thompson, 1991; Lev-Ari & Keysar, 2010; Livingston, Schilpzand, & Erez, 2017), less educated (Fraser & Kelly, 2012), less intelligent (Anderson et al., 2007; Fuertes, Potere, & Ramirez, 2002), and less hirable (Huang, Frideger, & Pearce, 2014). In this work, we found a possible advantage of being a nonnative speaker: Native speakers give foreign-accented speakers the benefit of the doubt when interpreting their utterances; as a result, implausible utterances are more likely to be interpreted in a plausible way when delivered in a foreign accent than in a native accent.

Recent work has demonstrated that when people understand language, they combine information about what is likely to be communicated—prior semantic expectations, or *priors*—with information on how messages can get corrupted by noise (Gibson, Bergen, & Piantadosi, 2013; Levy, 2008; Levy, Bicknell, Slattery, & Rayner, 2009). Gibson, Bergen, and Piantadosi formalized this account in terms of the following ideal-observer model (Geisler, 1989; Marr, 1982) of language comprehension, in which the comprehender engages in Bayesian decoding of the intended meaning:

$$p(s_i \mid s_p) \propto p(s_i) \times p(s_i \rightarrow s_p),$$

where s_p is the sentence perceived by the comprehender, and s_i is the sentence intended by the producer. The

Corresponding Author:
Edward Gibson, Massachusetts Institute of Technology, Department of Brain and Cognitive Sciences, 43 Vassar St., Building 46, Room 3035, Cambridge, MA 02139
E-mail: egibson@mit.edu

left-hand side of the proportionality, $p(s_i|s_p)$, gives the probability assigned by the comprehender to any particular hypothesized s_i given the observed linguistic input s_p. By Bayes's rule, this can be rewritten as the right-hand side of the proportionality, as a product of the prior probability $p(s_i)$ that a producer would wish to communicate s_i and the likelihood of the comprehender receiving s_p given that the speaker intended s_i, which is often notated as $p(s_p|s_i)$. This likelihood is written as $p(s_i \rightarrow s_p)$ to make it clear that it represents the probability of s_i being corrupted to s_p in the process of communication. The prior $p(s_i)$ represents all of the comprehender's relevant linguistic and world knowledge, including, for instance, the base rates of different grammatical constructions and the plausibility of different meanings. This term biases comprehenders toward utterances that are a priori plausible—things that are likely to be uttered. By trading off between the prior $p(s_i)$ and the likelihood $p(s_i \rightarrow s_p)$, comprehenders may arrive at interpretations that differ from the literal meanings of the specific sentences they perceive.

For example, consider the syntactic alternation between a sentence with a double object (DO) and a corresponding sentence with a single noun-phrase object plus a prepositional-phrase object (denoted PO in the literature):

a. DO, plausible: The mother gave the daughter the candle.

b. PO, plausible: The mother gave the candle to the daughter.

c. DO, implausible: The mother gave the candle the daughter.

d. PO, implausible: The mother gave the daughter to the candle.

Comprehension question: Did the daughter receive something/someone?

For the plausible versions, (a) and (b), *the candle* is the patient that is given to *the daughter*, and thus a literal reader should answer "yes" to the comprehension question. However, in the implausible versions, (c) and (d), the syntax suggests that *the daughter* is the patient that is given to *the candle*, a highly implausible event. Thus, a reader who relies on the literal meaning suggested by the syntax would answer "no" to the comprehension question, whereas a reader who infers that noise has somehow distorted the ordering or inclusion of words in (c) and (d) would answer "yes."

Notably, Gibson, Bergen, and Piantadosi (2013) further showed that comprehenders are highly sensitive to the overall amount of noise in the signal: As the perceived noise rate increases, participants rely more on their semantic priors (see also Gibson, Sandberg, Fedorenko, Bergen, & Kiran, 2016; Poppels & Levy, 2016). In Gibson, Bergen, and Piantadosi, the perceived noise rate was varied by manipulating the number of errors (misspellings, added or deleted words, word swaps, etc.) in the distractor materials. Thus, participants who encountered implausible statements, such as the ones in (c) and (d), in the context of many other sentences containing errors were more likely to infer the more plausible meaning (e.g., of the mother giving the candle to the daughter).

Communicating with nonnative speakers may lead to high perceived noise rates through a combination of (a) a priori knowledge that nonnative speakers are more likely to make syntactic errors and (b) situation-specific learning of the noise rate of the particular individual one is communicating with. Consistent with this framework, prior work suggests that comprehenders adapt to the higher error rate of nonnative speech. For example, in an event-related potential study, Hanulíková, van Alphen, van Goch, and Weber (2012) showed that the signature of the P600 component (a positive-going waveform that peaks around 600 ms after stimulus onset) is reduced for syntactic errors in accented speech. If the P600 indexes aspects of correcting errors (Gibson, Stearns, Bergen, Eddy, & Fedorenko, 2013), then Hanulíková et al.'s result suggests that listeners are more likely to assume that accented speech contains errors as a baseline, with the consequence that some errors are not corrected. An eye-tracking reading experiment conducted by Konieczny, Scheepers, and Hemforth (1994) in German also confirmed the predictions of the noisy-channel framework, although it was not originally discussed in these terms. Each participant in Konieczny et al.'s experiment was directed by one of two experimenters to read the sentences that would be presented to them. One of the experimenters was a native German speaker, and the other was a native English speaker who spoke German with an accent. There were noun-verb-noun sequences in the reading materials, which were disambiguated as object-verb-subject by German morphology, but which had equal plausibility as the (ungrammatical) subject-verb-object interpretation. Responses to questions following sentence trials indicated that the participants who were given the instructions by the nonnative German experimenter interpreted many of these noun-verb-noun sequences as subject-verb-object sequences—a much more frequent syntactic frame than object-verb-subject—whereas the participants who were given the instructions by the native German experimenter were more likely to interpret them as object-verb-subject sequences. Konieczny et al.'s results therefore suggest that when participants are interacting with nonnative speakers, they are likely to rely more on their syntactic prior for a noun-verb-noun sequence—the subject-verb-object interpretation—probably because they think that nonnative speakers might not know the rare but possible object-verb-subject interpretation.

In the current research, we tested whether the interpretation of sentences using strong world-knowledge biases is affected by the accent (or lack thereof) of the speaker. If so, this could create situations in which speakers with a nonnative accent have an advantage over speakers with a native accent, in that listeners might be more likely to interpret their implausible utterances in a more plausible way. That is, a straightforward prediction of the noisy-channel approach is that when communicating with nonnative speakers, comprehenders should give the speaker the benefit of the doubt and therefore be more likely to rely on their semantic priors in interpreting their utterances and less likely to interpret utterances strictly literally. We tested this prediction in a language-comprehension study consisting of six experiments in which participants listened to auditory versions of Gibson, Bergen, and Piantadosi's (2013) implausible materials, such as the ones in Examples (c) and (d) given previously, and answered simple comprehension questions. The critical manipulation in each experiment was whether the speaker had a heavy accent or no accent. The first three experiments investigated three syntactic alternations from Gibson, Bergen, and Piantadosi. The last three experiments were replications of the first three, with an additional control condition.

Experiments 1 Through 3

Method

Participants. We posted surveys for 960 distinct workers in three experiments, in each of which the sample was divided into four groups of 80 workers each. The surveys were posted on Amazon's Mechanical Turk, and the experiments were run using the Turkolizer software (Gibson, Piantadosi, & Fedorenko, 2011). Because Experiment 2 from Gibson, Bergen, and Piantadosi (2013) found an effect of increased perceived noise rate in 300 participants, we decided on a similar sample size (320 participants) for each experiment here.[1] All participants were paid. Participants were asked to indicate their native language and country of origin, but payment was not contingent on their responses to these questions. To constrain the population to American English speakers, we restricted the Internet protocol (IP) addresses to those in the United States. Furthermore, we filtered out participants who indicated that their native language was not English or that they were not originally from the United States. Following these restrictions, we removed 75 participants' data across the three experiments. In addition, we analyzed data only from participants who answered only one survey (they were instructed to fill out only one, but occasionally a participant filled out two or more) and who answered at least 75% of the questions for the 60 filler sentences correctly (the mean across participants and experiments was over 90% before excluding these participants). We removed a further 45 participants' data across the three experiments because of these restrictions, which left 840 participants across experiments for analyses (an average of 280 participants per experiment, corresponding to an average of 70 participants per list; see Procedure for information about the lists).

Design and materials. Each of the three between-participants experiments used one of three syntactic alternations, each crossing accent (accent, no accent) with the speaker of the materials (Speaker 1, Speaker 2). To counterbalance the identity of the speaker, we ensured that each set of target items was produced by two speakers: Speaker 1 (Idan Blank, from Israel) spoke the materials in near-native English with a strong Israeli accent, and Speaker 2 (Nezar Abdenur, from Canada, but with expertise as an actor speaking in many accents) spoke the materials in native English with a strong Hindi accent. These target materials were combined with the filler materials produced by both speakers recorded with no accent, for a total of four versions of each experiment: In two versions, Speaker 1 presented the target items (with and without accent, respectively), and Speaker 2 always provided the filler items; the other two versions had the same design, except that Speaker 2 presented the target items and Speaker 1 the filler items. Thus, any difference that we observed could not be due to the particular speaker.

Each of the three experiments used one of three syntactic alternations (from Gibson, Bergen, & Piantadosi, 2013). In Experiment 1, we used the DO-PO alternation, as in the example given in the introduction. In Experiment 2, the sentences differed according to whether the verbs were transitive or intransitive, as in the following example:

a. Transitive, plausible: The tax law benefited the businessman.
b. Intransitive, plausible: The businessman benefited from the tax law.
c. Transitive, implausible: The businessman benefited the tax law.
d. Intransitive, implausible: The tax law benefited from the businessman.

In Experiment 3, the sentences differed according to whether they were constructed in the active or passive voice, as the following example shows:

a. Active, plausible: The girl kicked the ball.
b. Passive, plausible: The ball was kicked by the girl.
c. Active, implausible: The ball kicked the girl.
d. Passive, implausible: The girl was kicked by the ball.

Crucially, only the implausible versions were used in Experiments 1 through 3. For the implausible materials, Gibson, Bergen, and Piantadosi (2013) found that participants made more inferences to the plausible interpretation of the DO-PO and transitive-intransitive alternations, compared with the active-passive alternation. In contrast, even in the presence of noise (errors) in the filler materials, participants interpreted the implausible active and passive constructions literally most of the time. Gibson, Bergen, and Piantadosi hypothesized that this difference between the DO-PO and transitive-intransitive constructions on the one hand and the active-passive constructions on the other was due to the noise likelihood. In particular, only one edit (addition or deletion of a function word) was needed to get from an implausible DO, PO, transitive, or intransitive construction to a more plausible alternative, whereas two edits were needed to get from an implausible active or passive construction to a more plausible alternative. Because fewer edits were required, a noise process was more likely to corrupt a plausible DO or PO utterance to an implausible one than it was to corrupt a plausible active or passive utterance to an implausible one. If our noise manipulation (accent vs. no accent) followed the pattern of Gibson, Bergen, and Piantadosi's results, we expected to find an effect of accent for the DO-PO and transitive-intransitive constructions but not for the active-passive constructions.

We used the 20 items from each of these three alternations from Gibson, Bergen, and Piantadosi's (2013) materials, along with their 60 filler items, to generate our auditory materials. The two speakers also produced the 60 filler items in native and near-native English. All materials are available at the Open Science Framework (https://osf.io/7c9bw).

Procedure. For each of the four versions of each experiment, we created two experimental lists. Each list contained the fillers and half of the target items, which were distributed between the lists following a Latin-square design. Each participant received one list, and the order of trials was randomized for each participant. All participants then read the following instructions: "This is a set of 80 auditory sentences. Answer the questions immediately following, according to what you think the speaker intended." There was a single yes/no question following each item (e.g., "Did the daughter receive something/someone?" "Did the girl kick something/someone?" "Did the tax law benefit from anything?"). Participants' answers to the questions following the target materials provided strong cues as to whether they interpreted the sentences literally (implausibly) or inferred the more plausible meaning. It took approximately 10 to 15 min for each participant to complete the task.

Evaluating the comprehensibility of the materials. Our critical measure was how often participants interpreted implausible items as their corresponding more plausible alternatives in the accent condition than in the no-accent condition. Our hypothesis was that participants would make a high-level inference about the likely meaning on the basis of the rate of noise in nonnative versus native speech. However, there could be a higher rate of plausibility-based interpretations in the accent condition for a less interesting reason: Perhaps participants simply would not be able to discern the words in the utterance and would answer at random or rely on the plausibility of the event on the basis of the partial information in the question. To test whether participants could accurately perceive the content of our accent materials, we performed a norming experiment.

An additional 480 Mechanical Turk participants were asked to transcribe what the speakers said, even if it was implausible. There were four surveys, one for each combination of speaker and accent. Each survey was given to 120 participants. On each of the four surveys, the 60 target items (20 items with two versions each, as in c and d in the examples given previously × 3 syntactic alternations) were divided across two lists, so that each participant made 60 transcriptions. Because we wanted to match the information that these participants got and the information that the participants in the critical experiments would receive, we presented the target sentences alongside the accompanying questions.

The transcriptions were coded for differences from the intended sentence in two ways: (a) whether a content word was misheard (e.g., "was in" instead of "worsened"; "boy" instead of "ball") and (b) whether a function word was added or deleted in order to arrive at a more plausible alternative (e.g., "The mother gave the candle to the daughter" instead of "The mother gave the candle the daughter"). A small number of recordings proved difficult for participants to understand, which led to inaccurate transcriptions on more than 50% of trials. For example, one speaker's accented version of "worsened" was transcribed as "was in" by more than half of the participants. These recordings (a total of six combinations of item, condition, speaker, and accent out of the 480 total recordings) were omitted from later analyses.[2] Furthermore, because we were most interested in inferences that participants made when hearing implausible sentences (not ones they misheard initially), we also omitted from later analyses combinations of item, condition, speaker, and accent for which participants made errors in their function-word transcriptions on more than 20% of the trials. This resulted in 6 further recordings being omitted, which left 468 (97.5%) of recordings to be analyzed in the critical experiments. The transcription error rates

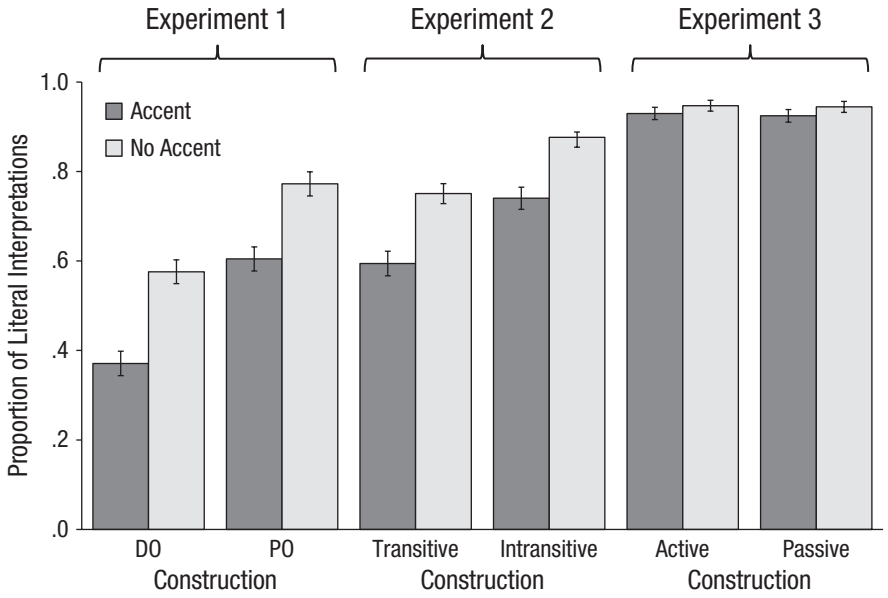

Fig. 1. Results from Experiments 1 through 3: mean proportion of literal interpretations of sentences as a function of sentence construction and whether or not the speaker uttered the sentences with a heavy accent. Lower values indicate more plausible interpretations. Results are collapsed across speakers. Error bars show 95% confidence intervals. DO = double object; PO = single noun-phrase object plus a prepositional-phrase object.

across conditions (see Table S1 in the Supplemental Material available online) were below 2% for all but the DO materials, and these had an error rate of only 3.4% (no-accent condition) and 5.9% (accent condition). Thus, the inference rates observed in the critical experiments for the DO-PO and transitive-intransitive constructions, which were between 12.9% (intransitive, no-accent condition) and 64.2% (DO, accent condition) cannot be explained by difficulties with discerning the utterance.

Results

Participants correctly answered comprehension questions for the filler sentences at a mean rate of 93% across experiments, varying between 91% and 96%, which suggests that participants were performing the required task. Our critical measure was how often participants interpreted implausible items as their corresponding more plausible alternatives in the accent than in the no-accent condition. The means and confidence intervals for each combination of speaker accent and implausible-sentence construction in all three experiments are presented in Figure 1 collapsed across speakers and in Table 1 separately for each speaker.

We analyzed the experiments using sum-coded mixed-effect logistic regressions (Gelman & Hill, 2007) with intercepts for participants and items, as well as slopes for accent (accent, no accent) and construction (e.g., DO-PO, transitive-intransitive, active-passive) for both participants and items in the random-effects structure. The active-passive experiment (Experiment 3) did not converge with slopes in the random-effects structure, but none of the critical main effects were close to significant in any analysis that we tried. This is probably because the means were close to ceiling in this experiment. Each experiment consisted of eight subexperiments: 2 constructions (e.g., DO, PO) × 2 accent conditions (accent, no accent) × 2 speakers (Speaker 1, Speaker 2).

There was a reliable main effect of speaker in Experiments 1 and 2; participants made more plausibility-based inferences for Speaker 1 than for Speaker 2 (Experiment 1: $\beta = 0.71$, $p = .01$; Experiment 2: $\beta = 0.77$, $p = .0007$). This effect was nonsignificant in Experiment 3. These main effects may simply mean that Speaker 1 had a stronger accent than Speaker 2, which may have led to a greater perceived noise rate.

As predicted by the noisy-channel hypothesis, the rate of literal interpretation was lower for the accent than for the no-accent conditions in Experiment 1 ($\beta = 1.41$, $p < .0001$) and Experiment 2 ($\beta = 1.16$, $p < .0001$). There was also an effect of construction in Experiment 1 ($\beta = 1.65$, $p < .0001$); people made more plausibility-based inferences for the DO than the PO construction. There was a similar effect of construction in Experiment 2 ($\beta = 1.27$, $p = .0005$), in which people made more plausibility-based inferences for the transitive than the intransitive construction. These within-experiments between-constructions differences replicated the results of Gibson, Bergen, and

Table 1. Results From Experiments 1 Through 3: Mean Proportion of Literal Interpretations of Sentences, Separately for Each Speaker

	Speaker	
Condition and accent	Idan	Nezar
	Experiment 1	
Double object		
Accent	0.304 [0.266, 0.343]	0.428 [0.390, 0.466]
No accent	0.529 [0.489, 0.569]	0.615 [0.580, 0.651]
Single object plus prepositional-phrase object		
Accent	0.533 [0.493, 0.574]	0.666 [0.631, 0.702]
No accent	0.755 [0.721, 0.789]	0.788 [0.758, 0.818]
	Experiment 2	
Transitive		
Accent	0.566 [0.527, 0.605]	0.623 [0.584, 0.661]
No accent	0.693 [0.660, 0.727]	0.808 [0.780, 0.837]
Intransitive		
Accent	0.705 [0.669, 0.741]	0.779 [0.745, 0.812]
No accent	0.832 [0.804, 0.859]	0.922 [0.902, 0.941]
	Experiment 3	
Active		
Accent	0.906 [0.884, 0.928]	0.955 [0.939, 0.970]
No accent	0.932 [0.913, 0.951]	0.964 [0.949, 0.979]
Passive		
Accent	0.897 [0.874, 0.920]	0.953 [0.937, 0.969]
No accent	0.942 [0.925, 0.960]	0.947 [0.929, 0.964]

Note: Values in brackets are 95% confidence intervals.

Piantadosi (2013), who argued that people make more plausibility-based inferences when the implausible version could be generated from the plausible alternative via deletion of a function word rather than via insertion of one. In Experiment 3, we found that accent had no reliable effect ($\beta = 0.85$, $p = .12$). This result is parallel to results from Gibson, Bergen, and Piantadosi, who found no effect on perceived noise rate in active-passive constructions, as manipulated by the inclusion of errors in the filler materials. Finally, we found no difference in inference rate between the active and passive constructions ($\beta = 0.29$, $p = .19$), which also replicated the results from Gibson, Bergen, and Piantadosi. There were no reliable interactions in any of the models. (See Tables S2, S3, and S4 in the Supplemental Material for full results of the models.)

Experiments 4 Through 6: Replications

Our effects might have been driven in part by the lack of plausible target materials spoken by the accented speaker.[3] Consequently, we ran a replication of all three experiments with plausible control materials spoken by the target speaker. That is, the designs of these three experiments were identical to those of Experiments 1 through 3, respectively, except that each set of target materials had four conditions: the two implausible conditions, and two plausible ones, as in (a) and (b) in the three examples given previously.

Method

Participants. Following the same procedures as in Experiments 1 through 3, we posted surveys for 960 workers on Mechanical Turk, all distinct from participants who took part in Experiments 1 through 3. Filtering out participants who indicated that their native language was not English or that they were not originally from the United States resulted in the elimination of 69 participants' data across the three experiments. In addition, analyzing data only from participants who answered at least 75% of the questions for the 60 filler sentences correctly (the mean across participants and experiments was over 90% before excluding these participants) resulted in the elimination of a further 84 participants' data across the three experiments, which left 807 participants across experiments for analyses (an average of 269 participants per experiment, corresponding to an average of 67 participants per list).

Design, materials, and procedure. As in Experiments 1 through 3, each experiment crossed the speaker of the

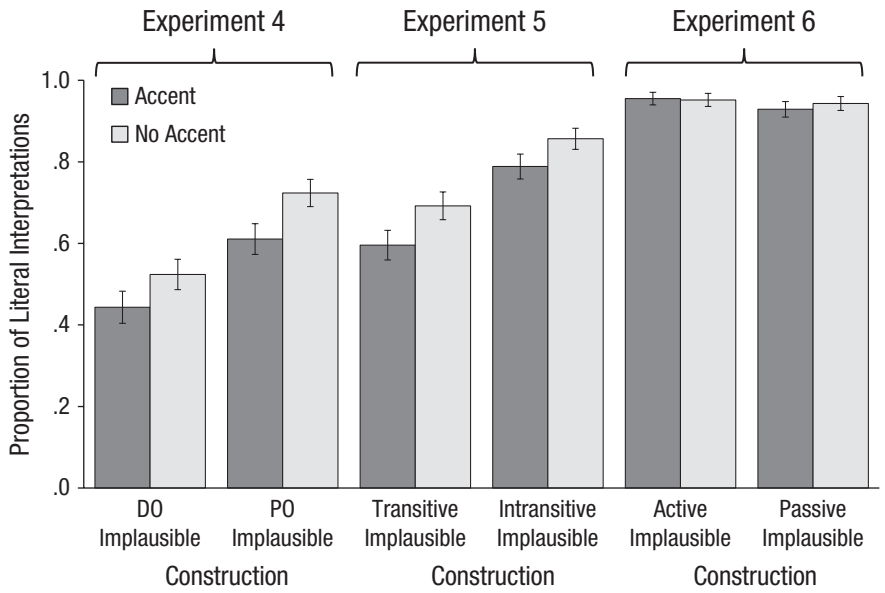

Fig. 2. Results from Experiments 4 through 6: mean proportion of literal interpretations of implausible sentences as a function of sentence construction and whether or not the speaker uttered the sentences with a heavy accent. Lower values indicate more plausible interpretations. Results are collapsed across speakers. Error bars show 95% confidence intervals. DO = double object; PO = single noun-phrase object plus a prepositional-phrase object.

materials (Speaker 1, Speaker 2) with whether he spoke with an accent (accent, no accent). Experiments 4 through 6 corresponded to the three syntactic alternations in Experiments 1 through 3, respectively. Unlike Experiments 1 through 3, both plausible and implausible versions were used in Experiments 4 through 6. We used the same 20 items from each of Experiments 1 through 3 (along with their 60 filler items), but broke them into a 2 × 2 design, crossing construction and plausibility. The accent and no-accent versions of one speaker's target materials were combined with the 60 filler items spoken by the other speaker. The procedure of Experiments 4 through 6 was identical to that of Experiments 1 through 3.

Results

Participants correctly answered comprehension questions for the filler sentences at a mean rate of 93% across experiments, varying between 92% and 94%, which suggests that participants were performing the required task. The means and confidence intervals for each combination of speaker accent and implausible-sentence construction are presented in Figure 2 collapsed across speakers and in Table 2 separately for each speaker.

The results of these replications were very similar to the results of Experiments 1 through 3. As before, we analyzed the three experiments using sum-coded mixed-effect logistic regressions (Gelman & Hill, 2007), with intercepts for both participants and items, as well as slopes for accent (accent, no accent) and construction (e.g., DO-PO, transitive-intransitive, active-passive) for both participants and items in the random-effects structure for each model. Each experiment consisted of eight subexperiments: 2 constructions (e.g., DO vs. PO) × 2 accent conditions (accent, no accent) × 2 speakers (Speaker 1, Speaker 2).

Because response accuracies on the plausible versions were near ceiling, we constrained our analyses to the implausible materials. This made the analyses parallel to those for Experiments 1 through 3. As predicted by the noisy-channel hypothesis, the rate of literal interpretation was lower for the accent than the no-accent conditions in Experiment 4 ($\beta = 0.68$, $p = .008$) and Experiment 5 ($\beta = 0.60$, $p = .002$). There was also an effect of construction in Experiment 4 ($\beta = 1.31$, $p < .0001$); people made more plausibility-based inferences for the DO than the PO construction. The same effect occurred in Experiment 5 ($\beta = 1.28$, $p = .0003$); people made more plausibility-based inferences for the transitive than for the intransitive construction. These within-experiments between-constructions differences (DO vs. PO; transitive vs. intransitive) again replicated the results of Gibson, Bergen, and Piantadosi (2013). In Experiment 6, we found that accent had no reliable effect ($\beta = 0.17$, $p = .68$). This result is also parallel to results from Gibson, Bergen, and Piantadosi. Finally, we found a small difference in inference rate between the active and passive constructions ($\beta = 0.46$, $p = .03$); people made more inferences for the active than passive constructions. This is in the direction predicted by the hypothesis that people make more plausibility-based inferences when an

Table 2. Results From Experiments 4 Through 6: Mean Proportion of Literal Interpretations of Implausible Sentences, Separately for Each Speaker

	Speaker	
Condition and accent	Idan	Nezar
	Experiment 4	
Double object		
Accent	0.390 [0.333, 0.447]	0.489 [0.435, 0.543]
No accent	0.514 [0.459, 0.569]	0.532 [0.482, 0.583]
Single object plus prepositional-phrase object		
Accent	0.579 [0.523, 0.635]	0.638 [0.587, 0.688]
No accent	0.710 [0.661, 0.760]	0.735 [0.690, 0.780]
	Experiment 5	
Transitive		
Accent	0.540 [0.487, 0.594]	0.647 [0.598, 0.696]
No accent	0.689 [0.641, 0.737]	0.696 [0.647, 0.744]
Intransitive		
Accent	0.797 [0.754, 0.840]	0.780 [0.737, 0.824]
No accent	0.856 [0.819, 0.892]	0.858 [0.821, 0.895]
	Experiment 6	
Active		
Accent	0.943 [0.919, 0.968]	0.966 [0.947, 0.985]
No accent	0.942 [0.917, 0.967]	0.962 [0.942, 0.982]
Passive		
Accent	0.922 [0.894, 0.951]	0.935 [0.910, 0.961]
No accent	0.919 [0.890, 0.948]	0.967 [0.948, 0.985]

Note: Values in brackets are 95% confidence intervals.

implausible sentence could be generated from a plausible alternative via deletion of a function word rather than via insertion of one, but this particular result is hard to interpret given the proximity of both conditions to ceiling (95% vs. 93% literal interpretations). There were no reliable interactions in any of the models. (See Tables S5, S6, and S7 in the Supplemental Material for full results of the models.)

Discussion

Inspired by a recent reconceptualization of high-level language interpretation as a combination of knowledge of (a) what is likely to be communicated (priors) and (b) how messages can get corrupted by noise during communication (e.g., Gibson, Bergen, and Piantadosi, 2013; Levy, 2008; Levy et al., 2009), we here examined the processing of accented speech. Gibson, Bergen, and Piantadosi previously showed that a greater perceived noise rate in the linguistic input (by adding errors) led comprehenders to rely more strongly on their semantic priors. We tested whether a similar increase in plausibility-based inferences would occur for accented speech. Indeed, across four constructions, we observed more plausibility-based inferences (~10%) for sentences produced with an accent than for sentences produced without an accent. Furthermore, we also showed that participants could correctly transcribe exactly what was spoken almost all of the time. Thus, our results suggest that, under certain circumstances, people may be more likely to give a nonnative speaker the benefit of the doubt when interpreting their utterances: People will assume that speakers with a foreign accent have more knowledge relative to what they literally say than a nonaccented speaker does.

It is an open question whether all accents are equally likely to induce plausibility-based inferences such as the ones discussed here. In our experiments, there were only two speakers: one who spoke English natively and could speak English with a Hindi accent well, and a near-native speaker of English who could speak English with an Israeli accent. There was no main effect of speaker in our experiments: Listeners made approximately the same inferences for each speaker. But it is possible that listeners would make more or fewer plausibility-based inferences depending on their sociolinguistic perception of the speaker, relative to their dialect of English. Future work should investigate these sociolinguistic consequences, varying both the target language (English in the current case) and the accented languages.

How can our results be reconciled with the observations that foreign-accented speakers are often attributed less credibility, intelligence, and education than native speakers? In terms of meaning interpretation, when one produces an incorrect or implausible sentence, there appears to be an advantage of being a nonnative compared with a native speaker because the utterance will be reinterpreted. But this also comes with the disadvantage of being perceived as syntactically unstable—native speakers also expect the nonnative speaker to make syntactic errors and are not very surprised by them (Hanulíková et al., 2012), which in turn might give rise to the perception of reduced intelligence and credibility.

But still, while previous work has shown disadvantages for speakers with foreign accents, the results reported here suggest an advantage to having a foreign accent in a particular situation. Imagine you want to appear knowledgeable about a topic but are in fact uncertain about it, perhaps at a cocktail party at which you want to make business connections. If you say something implausible or wrong, the person you are talking with may think less of you for your confusion. It would be advantageous for you if your implausible statement were interpreted as a more plausible, similar-sounding alternative. We demonstrated that such favorable misinterpretations are much more likely for nonnative than for native speakers. In the words of Arianna Huffington (2013), "I moved to New York in 1980 and met Henry Kissinger, who told me not to worry about my accent, because you can never, in American public life, underestimate the advantages of complete and total incomprehensibility."

Action Editor

Matthew A. Goldrick served as action editor for this article.

Author Contributions

E. Fedorenko, E. Gibson, L. Konieczny, and B. Hemforth developed the study concept. E. Gibson, C. Tan, R. Futrell, K. Mahowald, and E. Fedorenko designed the study. Testing and data collection were performed by E. Gibson. E. Gibson, R. Futrell, K. Mahowald, and L. Konieczny analyzed and interpreted the data. E. Gibson drafted the manuscript, and E. Fedorenko, R. Futrell, and L. Konieczny provided critical revisions. All authors approved the final version of the manuscript for submission.

Acknowledgments

We thank Idan Blank and Nezar Abdenur for producing the materials; Caitlyn Hoeflin for recording the materials; and Roger Levy, Alex Paunov, and Rachel Ryskin for their helpful comments.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

The work was supported by grants from the National Science Foundation Linguistics Program (Award 1534318 to E. Gibson), the Simons Center for the Social Brain at the Massachusetts Institute of Technology, the Brain and Language Research Institute (No. ANR-11-LABX-0036), and A*MIDEX (No. ANR-11-IDEX-0001-02).

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617690277

Open Practices

All data and materials have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/7c9bw/. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617690277. This article has received the badges for Open Data and Open Materials. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

Notes

1. Furthermore, we used 320 participants in the replication experiments in the present study (Experiments 4–6) and also found robust effects.
2. The results were qualitatively the same with or without these omissions. All relevant inferential statistics resulted in similar effect sizes and similar p values in significance tests. Furthermore, the results were qualitatively the same when more combinations of item, condition, speaker, and accent were filtered (with the corresponding effect that differences in error rates on remaining data were almost eliminated). For example, if we filtered combinations of item, condition, speaker, and accent with error rates of 10% or more, then all relevant inferential statistics still resulted in similar effect sizes and similar p values in significance tests.
3. We thank a reviewer (Kristin Lemhöfer) for suggesting this possibility.

References

Anderson, A., Downs, S. D., Faucette, K., Griffin, J., King, T., & Woolstenhulme, S. (2007). How accents affect perception of intelligence, physical attractiveness, and trustworthiness of Middle-Eastern-, Latin-American-, British-, and Standard-American-English-accented speakers. *BYU Undergraduate Journal of Psychology*, *3*, 5–11.

Bourdieu, P., & Thompson, J. B. (1991). *Language and symbolic power*. Cambridge, MA: Harvard University Press.

Fraser, C., & Kelly, B. F. (2012). Listening between the lines: Social assumptions around foreign accents. *Australian Review of Applied Linguistics*, *35*, 74–93.

Fuertes, J. N., Potere, J. C., & Ramirez, K. Y. (2002). Effects of speech accents on interpersonal evaluations: Implications

for counseling practice and research. *Cultural Diversity & Ethnic Minority Psychology*, *8*, 346–356.

Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review*, *96*, 267–314.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models* (Vol. 1). New York, NY: Cambridge University Press.

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences, USA*, *110*, 8051–8056.

Gibson, E., Piantadosi, S., & Fedorenko, K. (2011). Using Mechanical Turk to obtain and analyze English acceptability judgments. *Language and Linguistics Compass*, *5*, 509–524.

Gibson, E., Sandberg, C., Fedorenko, E., Bergen, L., & Kiran, S. (2016). A rational inference approach to aphasic language comprehension. *Aphasiology*, *30*, 1341–1360.

Gibson, E., Stearns, L., Bergen, L., Eddy, M., & Fedorenko, E. (2013, March). *The P600 indexes rational error correction within a noisy-channel model of human communication*. Paper presented at the 26th Annual CUNY Conference on Human Sentence Processing, Columbia, SC.

Gluszek, A., & Dovidio, J. F. (2010). The way *they* speak: A social psychological perspective on the stigma of non-native accents in communication. *Personality and Social Psychology Review*, *14*, 214–237.

Hanulíková, A., van Alphen, P. M., van Goch, M. M., & Weber, A. (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience*, *24*, 878–887.

Huang, L., Frideger, M., & Pearce, J. L. (2014). How non-native speakers can crack the glass ceiling. *Harvard Business Review*, *92*(6), 27–28.

Huffington, A. (2013). Arianna Huffington's 2013 Smith College commencement address [Video]. Retrieved from https://www.youtube.com/watch?v=UJ25qEHgcM4

Konieczny, L., Scheepers, C., & Hemforth, B. (1994). Reanalyses vs. internal repairs: Non-monotonic processes in sentence perception. In B. Hemforth, L. Konieczny, C. Scheepers, & G. Strube (Eds.), *First analysis, reanalysis, and repair* (IIG-Berichte 8/94; pp. 1–21). Department of Telematics, Institute of Computer Science and Social Studies, University of Freiburg, Germany.

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, *46*, 1093–1096.

Levy, R. (2008). A noisy-channel model of rational human sentence comprehension under uncertain input. In *Proceedings of the 13th Conference on Empirical Methods in Natural Language Processing* (pp. 234–243). Stroudsburg, PA: Association for Computational Linguistics.

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences, USA*, *106*, 21086–21090.

Livingston, B. A., Schilpzand, P., & Erez, A. (2017). Not what you expected to hear: Accented messages and their effect on choice. *Journal of Management, 43,* 804–833. doi:10.1177/0149206314541151.

Marr, D. (1982). *Vision*. New York, NY: W. H. Freeman.

Poppels, T., & Levy, R. (2016). Structure-sensitive noise inference: Comprehenders expect exchange errors. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (pp. 378–383). Austin, TX: Cognitive Science Society.