Prosody and the meanings of English negative indefinites

Frances Blanchette

Marianna Nadeu

January 2018

Abstract

This paper investigates the acoustic correlates of single and Double Negation (DN) readings of English negative indefinites in question-answer pairs. Productions of four negative words (*no one*, *nobody*, *nothing*, and *nowhere*) were elicited from 20 native English speakers as responses to negative questions such as "What didn't you eat?" in contexts designed to generate either a single negation reading or a logically affirmative DN reading. A control condition with no negation in the question was employed for comparison. A verification question following each item determined whether tokens were interpreted as expected and, therefore, produced with the target interpretation. Statistical analysis of the f0 curves revealed a significant difference: DN is associated with a higher fundamental frequency than single negation. In contrast, the single negative and control conditions were not significantly different with respect to f0. Analysis of the verification question responses showed significant differences between all three conditions (Control > DN > single negation), suggesting that single negation is more difficult to interpret than DN as a response to a negative question. The results are compared with previous work on Romance, and we demonstrate how English behaves like a prototypical Negative Concord language in that DN is the prosodically marked form.

Keywords: Double Negation, Negative Concord, prosody, denial negation, Standard English

1. Introduction

Consider the following context: A professor assigns a student a lengthy and challenging set of readings, but later realizes the assignment may have been too difficult. They meet, and the dialogue in (1) ensues.

(1) Professor: So, what didn't you read?
Student: Nothing.

The negative indefinite *nothing* in (1) is ambiguous between a single and a double negation reading. On the double negation (DN) reading, *nothing* means 'everything': The student implies that, contrary to expectation, there is nothing she did not read. On the single negation reading, *nothing* simply means 'nothing': The student confirms the professor's expectation and implies that in fact she did none of the assigned reading.

Previous studies have shown that, under certain pragmatic and prosodic conditions, DN readings exist in the prototypical Negative Concord (NC) languages. Espinal & Prieto (2011) demonstrate how Catalan speakers reliably associate DN readings with negative words used as responses to negative questions as in (1) when pronounced with a contradictory intonation contour (Liberman & Sag 1974). Espinal et al. (2016) examine Spanish and Catalan, showing that manipulations in both syntax and prosody can independently and reliably yield DN interpretations in these languages, and Déprez and Yeaton (to appear) find similar results for French. Prieto et al. (2013) show how gesture works in conjunction with prosody to enhance the accessibility of DN readings in Spanish and Catalan.

The current study builds on previous experimental findings for prototypical NC languages to examine the acoustic correlates of DN readings in English of negative indefinites in negative question-answer pairs such as (1). This introduction lays the background for the two experiments we present in Sections 2 and 3, as well as for our results discussion in Section

4. In Section 1.2 we briefly review aspects of the syntax of negative question-answer pairs, and in 1.3 we introduce and discuss their pragmatics. Section 1.4 discusses the relationship between prosody and negative meanings, in English and crosslinguistically, and includes a brief discussion of the related phenomenon of polar particle answers to polar questions. Section 1.5 concludes the introduction by laying out the objectives for our two experiments. We begin first in 1.1 with a discussion of the related phenomenon of two negatives sentences in English, which have the unique property of being heavily conditioned by language-external prescriptive pressures and norms.

1.1 Negative Concord and Double Negation in English

In NC sentences, two or more syntactic negations mark a single semantic negation, as the following example and its prose translation show:

(2) The student didn't do none of the assigned reading.
'The student did none of the assigned reading.'

English is historically an NC language, but, in present day English, NC is heavily socially stigmatized. At the time of prescriptive grammarian Bishop Lowth's 1762 edict that, in English, two negatives should equal a positive (Horn 2010), the presence of NC in formal written texts had already diminished significantly (Nevalainen 2006).[1] Despite its proscription, English NC has persisted and has come to be associated with "non-Standard" varieties including Appalachian (Wolfram & Christian 1976) and African American English (Green

[1] See Nevalainen (2006) on the social motivations for this shift.

2002) in the United States, as well as varieties of British (Anderwald 2002, 2005; Tubau 2016), Scottish (Smith 2001), and Irish English (Henry 2016).

In contrast with these so-called "non-Standard" varieties, it is widely accepted that, synchronically, "Standard English" is a DN language, which does not have NC (Ladusaw 1992; Déprez 2000, 2011; Watanabe 2004; Zeijlstra 2004; Kallel 2007; De Swart 2010; Espinal & Prieto 2011; Wallage 2012; Puskás 2012; Prieto et al. 2013; Longobardi 2014; Déprez et al. 2015; Espinal et al. 2016; Thornton et al. 2016; Tubau 2016; and others). More generally, models of DN and NC fall under the category of macroparametric approaches, in which languages are either DN or NC (Zeijlstra 2004), or that of microparametric approaches (Déprez 2011), in which particular (morpho-)syntactic and pragmatic conditions yield DN and NC readings in a single language. Both types of approach have typically modeled Standard English as having a distinct grammatical system from "non-Standard" Englishes as well as other NC languages with respect to negation (e.g. Espinal & Tubau 2016; Ladusaw 1992; Tubau 2016; Zeijlstra 2004).

A growing body of work demonstrates that DN readings are possible in NC languages (Espinal & Prieto 2011; Prieto et al. 2013; Déprez et al. 2015; Espinal et al. 2016; Déprez & Yeaton, to appear).[2] These authors have demonstrated that speakers of prototypical NC languages reliably interpret both NC and DN given predictable combinations of prosodic and pragmatic features. In a similar spirit, Blanchette (2017) asks whether speakers of Standard English, thought to be a DN language, can be shown to have reliable intuitions about NC. She hypothesizes that in traditional forms of data such as binary acceptability judgments, the heavy social stigma associated with English NC may mask speakers' natural intuitions about its grammatical properties, and conducts a series of quantitative gradient acceptability judgment

[2] See also De Swart and Sag (2002) and Longobardi (2014).

studies aimed at uncovering those intuitions. Blanchette's results show that Standard English speakers display a clear syntactic preference for NC constructions with a negative object over those with a negative subject. In addition, her results demonstrate a clear preference for single negation (NC) contexts (3a) over DN contexts (3b) for constructions with a negative object following a negative marker, as in the following example:

(3) She didn't do no reading last night.

(a) NC context: Maria fell asleep before she could even start her assignments.

(b) DN context: Maria usually skips the reading assignments, but last night was different.

The result that participants preferred NC contexts for items like (3) contradicts theories that assume Standard English is strictly a DN language, and suggests that the negation system in this English variety is more similar to prototypical NC languages than previously thought.

In this study, we adapt the experimental paradigm in Espinal and Prieto (2011) (see also Espinal et al. (2016)) to investigate whether DN and single negation readings in Standard English behave similarly to Spanish and Catalan, using question-answer pairs like (1). Our experiment involves undergraduate students in a university laboratory setting, which we assume elicits the use of "Standard English" in the sense generally understood within the field of linguistics. As such, our study contributes information on the prosodic patterns associated with single and double negation readings of negative indefinites in negative question-answer pairs in Standard English, and the extent to which these patterns overlap with those found in prototypical NC languages.

1.2 The Syntax of Single and Double Negation in Negative Question-Answer Pairs

Standard syntactic models assume that the answer in question and answer pairs like (1) involves an elided structure in which the negative phrase has raised to the left periphery, and the remainder of the sentence has undergone deletion at the Phonological Form (PF) interface (Merchant 2001; Temmerman 2012). Under the ellipsis analysis, the structure for *nothing* in (1) would be roughly as in (4):

(4) $[_{CP}$ nothing$_1$ $[_{TP}$ ~~I didn't read~~ $t_1]]$

Under standard syntactic models of NC (cf. Zeijlstra 2004), if a concord relation occurs between the two negations, this relation is established within the Tense Phrase (TP) in (4), prior to quantifier raising.

Espinal and Tubau (2016) argue against ellipsis accounts of negative fragment answers to negative questions in "NC languages" on the grounds that they incorrectly predict the DN reading to be the only possible one (pp. 49, 53). These authors propose an alternative model, in which the two readings are generated for a subset of speakers who have two distinct lexical entries for the negative word. One of these lexical items is specified for an uninterpretable negative feature, and combines with the negation introduced by the question to generate the single negation reading. The other carries its own logical negation, which cancels the negation introduced by the question, yielding the DN reading.

Note that under Espinal and Tubau's (2016) account, the single negation reading of the negative fragment response to the negative question is not NC in the technical sense, in that there is no concord relation established between two negations within a single clause. Following previous work by these authors, we continue to use the term NC to refer to our experimental stimuli, with the understanding that this "NC" may not be of the same syntactic nature as the NC that occurs within a clause. The terminology also serves an expository purpose

in that it serves to distinguish the critical items, with a negation in the question and the answer, from our control items, which have only a single negation in the answer (described below in Section 2).

1.3 The Pragmatics of DN and single negation in Negative Question-Answer Pairs

Consider again the context in (1), which contains a syntactic negation in both the question (*What didn't you read?*), and the fragment answer (*nothing*). The structure of this brief discourse serves as the basis of our experimental design. To address the question of how the single negation and DN readings of *nothing* in (3) are inferred, we must first address the meaning of the negation-containing question. Assume first, following standard semantic models, that the denotation of a question includes the set of its possible answers that are true (Hamblin 1973; Kartunnen 1977).[3] Imagine now that the professor had assigned two monographs: Rizzi (1990) and Kayne (1994). The following then represents the question's denotation set in that it exhausts the possible answers to the question *What didn't you read?*:

(5a) I didn't read Rizzi.

(5b) I didn't read Kayne.

(5c) I didn't read Rizzi or Kayne.

Note that (5c) represents the single negation response: It is true in a world in which the student read *nothing*, or neither of the two assigned monographs.[4] Note further that the DN response,

[3] See Dayal (2016) for an extensive review and synthesis.

[4] See Dayal (2016:44,46) on plural responses to questions in which the *wh*-expression is not specified for number.

which in this context contributes the meaning that the student read both of the assigned manuscripts, is not present in the question's denotation set. How, then, is this reading derived? To answer this, we must consider the interaction between the negations in the question and the answer, which, following previous work, we analyze in terms of the pragmatic notion of denial (Espinal & Prieto 2011; Prieto et al. 2013; Espinal & Tubau 2016; Espinal et al. 2016).

Geurts (1998) proposes a typology with four types of denial negation, each of which is directed at a previous utterance, hence metalinguistic in the sense of Horn (1989[2001]).[5,6] Proposition denial is directed at an assertion introduced by a previous utterance. Under this definition, if the denotation set of the question *What didn't you read?* is as in (5a–c), then the DN reading of the response *nothing* can plausibly be categorized proposition denial, in that it denies the truth of all of the propositions in the question's denotation set. Note that under this analysis, it is only the DN reading, and not the single negation reading, which can be thought of as a denial negation. This is because, in our representation in (5), the question's denotation set includes a proposition that makes the single negation response true, namely (5c). In this sense, the DN reading is derived through interaction with the pragmatic context in a way that the single negation reading is not.

The single negation reading, however, is not immune to interacting with assumptions introduced by the question. To understand how, let us consider Abusch's (2010) notion of "soft presupposition projection. Abusch proposes that there is a set of construction types, including

[5] Geurts (1998:292) argues that they are not all "purely metalinguistic", in that they are directed at both linguistic and metalinguistic objects.

[6] Geurts (1998) analyzes both implicature and form denials in terms of semantic transfer, where form or quotation denial targets the way in which the previous utterance was stated, and implicature denial targets scalar expressions like 'Mary ate five cupcakes', where 'five' implies an exactly reading. Presupposition denial (the fourth type) is discussed further below in the text.

questions, which introduce presuppositions that can be characterized as "soft" or "defeasible" (p. 6). To illustrate, consider the following (her example (16b) on p. 6):

(6) I've alienated my colleagues completely. Who will vote for me? Probably nobody.

In (6), the question *Who will vote for me?* contributes an existence presupposition: It is associated with the intuition that there exists at least one colleague who will vote for the speaker. The fact that the question can be answered with *nobody* shows that the presupposition is defeasible, hence "soft": The negative word response asserts that there are no individuals who satisfy the existence presupposition, and the response is perfectly felicitous in this context.

Returning to negative questions, note that the question *what didn't you read* in the context in (1) has an intuitive reading in which some but not all the assignments were completed. This intuition can be thought of as a pair of existence presuppositions: (i) There exists some assigned material that was read, and (ii) there exists some assigned material that was not read. Because they are contributed by a question, under Abusch (2010) they are "soft", hence defeasible. Note that under the single negation reading (i.e., (5c) in the question's denotation set), there exists no material that was read, and the presupposition in (i) is in fact left unsatisfied. The soft, defeasible nature of this presupposition is therefore crucial to the single negation interpretation of the negative phrase in a negative question-answer pair.

With this notion of soft presupposition defeasibility in place, we return briefly to Geurts's (1998) notion of denial negation. Under the assumption that the negative question what *didn't you read?* contributes the presupposition in (ii), that there exists material that was not read, then the DN response can also be plausibly construed of as a case of Presupposition denial. Under its DN reading, the response *nothing* implies that, contrary to expectation, there exists no material that was not read, a denial that is targeted directly at the presupposition in

(ii). The notion of defeasibility is thus fundamentally distinct from that of denial. In uttering *nothing* with its single negation interpretation, the speaker is not intentionally targeting a presupposition contributed by the question. Instead, the presupposition simply ceases to be a factor in interpretation. By contrast, as a presupposition denial, the DN interpretation is reliant on the presupposition itself, and the negative word (and crucially, as we will show, its acoustic properties) serve to deny it directly.

The experimental data we report in Section 2 demonstrate how unlike single negation, the special pragmatic status of denial (whether at the presuppositional or propositional level) is marked prosodically on negative indefinite fragment answers in English. In anticipation of our discussion of the acoustic results, the next sub-section discusses the relationship between prosody and information structure, and their relationship to negation.

1.4 Prosody, Meaning, and Negation

Prosody serves to highlight certain words and to break larger units of speech into smaller units, facilitating speech perception and processing. It also serves to convey information about sentence type, the structure of utterances, the status of entities in the discourse, pragmatic meaning, and information about the context in which speech is produced (Cole 2015). Languages, however, differ in their uses of prosody. For example, English prosodically marks information that is new in the discourse, and deaccents information that is discourse-salient or given (anaphoric deaccenting), whereas other languages do not (Ladd 2008). Similarly, some languages impose different prosody to a sentence like "Jessica got a promotion" depending on whether it responds to the question "What happened?", which elicits broad focus, or "What did Jessica get?", which elicits narrow focus on the object (Gussenhoven 2007; Elordieta 2007), whereas other languages do not distinguish prosodically between these two types of focus.

Previous work has suggested that in English, negation is also expected to have prosodic prominence, because it presents new information (O'Shaugnessy & Allen 1983; Pierrehumbert & Hirschberg 1990). However, studies of negation in spontaneous speech have shown that this is not consistently so (Yaeger-Dror 1995, 1997; Kaufman 2002). Rather, it has been demonstrated that prosodic prominence on negation varies according to the specific discourse function it contributes (Kaufman 2002).

Section 1.3 established that because it serves as a denial, the DN reading is pragmatically marked in a way that the single negation reading is not. Previous work on the acoustic correlates of information structure in English demonstrates that focus, a pragmatically marked discourse function, and contrastive focus in particular, is marked by higher intensity, longer duration, and higher f0 (Breen et al. 2010). Because of its marked discourse function, on the basis of this previous work, English DN is predicted to be prosodically marked relative to single negation. The acoustic data we present in section 2 show precisely how this prediction is borne out for the case of negative fragment answers to negative questions. Next we briefly discuss research on the marking of DN crosslinguistically, to situate our English results in the broader context of natural language.

1.4.1 The prosodic markedness of DN across languages

The body of work in Espinal and Prieto (2011), Prieto et al. (2013), and Espinal et al. (2016) has demonstrated that, in Spanish and Catalan, DN interpretations of single negative words in question-answer pairs like (1) are characterized by an intonational form that is different than that of the single negation reading. These authors found that the DN interpretation of negative indefinites is reliably associated with a "contradiction intonation" contour (Prieto et al. 2013: 145,147). This pattern, represented as L+H* L!H% in the Autosegmental-Metrical (AM) model of intonational phonology (Pierrehumbert 1980; Beckman & Pierrehumbert 1986; Ladd 2008),

is characterized by rising f0 on the stressed syllable followed by a complex falling and rising pattern on the posttonic (Prieto et al. 2013: 140).[7] This contrasts with the rising falling pattern of the single negation interpretation (L+H* L%), the typical pattern for broad, non-contradictory focus (Prieto et al. 2013: 140). DN interpretations have also been shown to be prosodically marked in French, also widely considered to be a prototypical NC language. Déprez & Yeaton (to appear) show that native French speakers mark negative words in DN constructions with a slight but reliable increase in f0, followed by a significant decrease in pitch.

Though the work mentioned above exhausts the systematic and controlled experimental investigations of DN prosody, there seems to be consensus in the literature that prosody plays an important role in the interpretation of DN crosslinguistically. For example, Biberauer & Zeijlstra (2012:357) suggest that prosodic marking is a necessary condition for DN to be felicitous in varieties of Afrikaans. Alonso-Ovalle & Guerzoni (2004:3) describe a similar state of affairs for Italian, showing that prosody serves to mark the DN reading in contexts for metalinguistic negation (described as a "rise and fall" in Zanuttini 1991:130). Puskás (2012) describes a similar state of affairs for Hungarian, another prototypical NC language. On her description, DN readings in Hungarian are marked by a "fall-rise" intonation pattern on the negative word (p. 619). Though these studies do not provide the specific acoustic correlates or patterns of intonational phrasing, the intuitions are clear: In addition to being pragmatically

[7] In the AM model, this notation represents the nuclear configuration associated with DN readings. The nuclear configuration or intonational contour includes a pitch accent (an f0 movement anchored on the nuclear stressed syllable) and a boundary tone (an f0 movement associated with the end of an intonational unit). For an introduction to the AM model, see Pierrehumbert (1980), Beckman and Pierrehumbert (1986), or Ladd (2008).

conditioned, DN is the prosodically marked form relative to single negation, across languages and even language families.

1.4.2 Polar particle responses to polar questions

Goodhue and Wagner (2018) examine the production and interpretation of bare polar particles in response to negative sentences, as in contexts like the following (adapted from their (39) on p. 22):

(7) Context: A work meeting has started at 4, and your colleague comes to your door at 4:07 to ask if you're planning to come.
Question: Are you not coming to the meeting?
Response: No, I'm not coming to the meeting.

The response particle no in (7) serves to confirm the assumption introduced by the polar question.[8] In this sense, it behaves similar to the single negation interpretation of negative word responses to negative *wh*-questions, the object under study in this paper. Goodhue and Wagner examine both positive (*yes*) and negative particle responses in contexts that elicit congruent (as in (7)) and contradictory interpretations of the particle, the latter of which bear similarity to the DN use of negative words such as *nothing*. Their production study shows that the most common overall (> 70%) intonational pattern in their data is the declarative fall (H* L-L%). The second most common contour is the so-called contradiction contour, characterized by a fall rise

[8] We have oversimplified here for expository purposes. For more on polar questions and particle answers, see Holmberg (2016), Kramer & Rawlins (2009), Krifka (2013), Roelofsen & Farkas (2015), and Holmberg (2016), among others.

movement (%H L* L-H%). This contour is typically found in positive responses to negative questions.

Goodhue and Wagner (2018) also report the results of a perception study, which reveal that, like negative fragment answers to negative questions, the prosodic rendering of bare polar particles conditions the way they are interpreted. In particular, when *yes* is produced with the so-called contradiction contour, it is interpreted as a positive response 65% of the time, whereas responses are at chance level when it is produced with a falling intonation. In conjunction with their production task, this work shows that speakers reliably use prosodic means to communicate and interpret the pragmatic meaning particle answers to negative questions. Crucially, they show that in English, responses that contradict assumptions introduced by the question are marked relative to responses that are neutral with respect to the question's assumptions.

1.5 Objectives

In light of the heavy prescriptive pressures shaping English negation, we suggest that in addition to providing a more accurate description of the contribution of prosody to the interpretation of English negation, our acoustic data can serve as an informative supplement to acceptability judgment and usage data in the construction of theoretical models. In the spirit of Lewis & Phillips (2015), who discuss how data from offline judgments and online, timed measures provide distinct windows onto the same grammatical phenomena, we submit that our acoustic data can inform the question of whether Standard English allows the inference of single negation readings, and whether and how the DN reading is marked relative to the single negation.

The experiments we present here build on previous work to further understand the contribution of prosody in the expression and interpretation of negation in context. The studies

were conducted with university students in a laboratory setting, an environment that, we argue, elicits the use of "Standard English". Our aim is to investigate the prosodic strategies that this English variety employs to generate DN readings of single-word utterances. On the basis of previous work, we expect that the pragmatically marked status of denial will be instantiated by some form of prosodic "markedness" in comparison with single negative, non-denial interpretations (Geurts 1998; Breen et al. 2010). One possibility is that, like Catalan, English uses a specific intonational contour (a different type of pitch accent and boundary tone combination) to signal the DN reading of two syntactic negations. Another possibility is that DN renditions exhibit phonetic features associated with prosodic prominence (e.g. higher intensity and/or f0, or longer duration), while being realized with the same nuclear contour.

Additionally, we aim to investigate whether Standard English, thought to be prototypically DN, displays properties similar to or different from prototypical NC languages by comparing our results to previous work on these languages investigating similar questions.

2. Production Experiment

2.1 Participants

22 undergraduates were recruited from an introductory linguistics class at a public university in the United States. We discarded data from two participants, one male and one non-native speaker, leaving us with data from 20 female native American English speakers with ages ranging between 18 and 22 (M = 20.04, SD = 1.15).[9] Following the experiment, participants completed a language history questionnaire. Five different U.S. states were represented,

[9] Because students received credit for a course assignment in exchange for participation, for the sake of fairness, we tested non-native speakers even though they did not meet our criteria for participation. Since only one male signed up, and given that we did not have sufficient male voices to balance out the sample, we also discarded his data.

including Kansas (1), Kentucky (1), Massachusetts (1), New Jersey (3), and Pennsylvania (14). Three of the participants from Pennsylvania reported that they grew up in rural areas, and the remainder of the participants had suburban upbringings. Three reported having at least some knowledge of another language, 15 reported having had musical training, with experience ranging from 2 to 11 years and including both voice and instrumental, and none reported any hearing or speaking impairments (except for one speaker, who reported having been diagnosed with a lisp).

The university where this work was conducted is competitive and has relatively high academic standards. In order to be accepted, native English-speaking students must have previously demonstrated proficiency in Standard English through various standardized achievement tests and other formal prerequisites, and they are expected to employ this version of English in the classroom and when completing assignments (Johnson & VanBrackle 2011; Dunstan & Jaeger 2015; Horton 2017). We therefore assume that, in this formal environment, participants were primed to use their version of Standard English.

2.2 Materials

The experiment was designed to elicit productions of the words *nothing*, *nobody*, *no one*, and *nowhere* in three different conditions. Following Prieto et al. (2013), we provided contexts that elicited either an NC or a DN interpretation of the target word. Each item included a context and a question, followed by a single word response, which participants were instructed to read aloud. We instructed participants to imagine themselves in the context and to read the word as though they were in the context, adding emphasis where necessary.[10]

[10] Instructions were presented as follows: "In this experiment, we will ask you to imagine yourself in different contexts. You will read a context, then a question. The following screen will contain a word (or words) in green. Please say the words in green out loud. When you say the words, say them as if

Following the instructions, participants completed three practice items, and had the opportunity to ask questions about the protocol before beginning the experiment.

Test items were divided into three conditions: NC, DN, and Control. The following examples illustrate an NC and a DN item for the word *nothing*:

(8) Negative Concord

Context: You and your roommate pay different bills each month. This month you have too little money to pay bills.

Question: Your roommate asks: What didn't you pay?

Response: Nothing.

(9) Double Negation

Context: You and your roommate pay different bills each month. This month you surprise your roommate and pay all of the bills.

Question: Your roommate asks: What didn't you pay?

Response: Nothing.

---

you were really in the context. You should find that some of the contexts and questions require an emphatic answer. Please try to convey that emphasis when you say the word or words in green. Lastly, you will be asked a True or False question about each context. The button with the green sticker (d) is True, and the one with the red sticker (k) is False. Once you have moved on you cannot go back to a previous screen. For this reason, it is important to read and think about each context carefully. You will now have a chance to practice. Remember to say the word in green, and use emphasis if necessary."

In (8), participants are asked to imagine themselves in a scenario in which they cannot pay the bills. In this context, the roommate's question about what was not paid elicits a single negation interpretation of the fragment *nothing* (i.e., that nothing was paid). In (9), the context asserts that the participant paid all the bills, and that nothing was not paid, the denial reading.[11]

Our stimuli included ten NC and ten DN items, and each participant produced five tokens of each of the four negative words. As illustrated in (8) and (9), each NC question had a parallel DN counterpart (and vice versa). Participants were divided into two subgroups, and received either the NC or the DN context for each item (e.g., either (8) or (9)), but not both.

Eight control items (two for each negative word) with no negation in the question were also included to determine how participants pronounce the words when there is only one negation and the response is neither NC nor DN. The following illustrates a control item:

(10) Single Negative Control

Context: You are having dinner at your friend's house. You forgot to bring dessert.

Question: Your friend asks: What did you bring?

Response: Nothing.

[11] In items (8) and (9) the *wh*-phrase questions a direct object. English NC is sensitive to the position of the negative marker with respect to the negative phrase (Blanchette 2017). All English varieties in which NC is realized employ constructions with a negative direct object and preceding negative marker, but only a subset employ negative subjects in NC (Smith 2001; Anderwald 2002, 2005). Therefore, to establish a baseline for comparison and to avoid introducing potential confounds related to microvariation in English NC, the *wh*-phrases in our question-answer pairs questioned direct objects only.

Items for this experiment were interspersed with items for another experiment with a similar methodology, which served as distractors and are thus not discussed here.

After pronouncing each negative word, participants had to perform an additional task which allowed us to assess their comprehension of the word in context. Each item in the experiment was followed by verification question which included a statement paraphrasing its meaning that participants had to judge as true or false. For the critical items, half of the statements were true on the NC reading, and half were true on the DN reading. Both subgroups of participants received the same statements, but because the context types were reversed, the target answer was also reversed. For example, the statements for both items (8) and (9) was "You paid every household bill this month". For the group who received (8), which elicits an NC interpretation, the statement was false, but for the group that received (4), which elicits a DN interpretation, the statement was true.

Inclusion of the verification question allowed us to assess whether participants interpreted the item as intended. This in turn allowed us to exclude non-target renditions from the acoustic analysis. In addition, these data constituted a behavioral measure that allowed us to compare participants' comprehension of the single and double negation contexts and items.

2.3 Procedure

Each participant saw a total of 48 items in context, including 20 critical items, 8 controls, and 20 distractors. The stimuli were presented electronically using E-Prime 2 Software (Psychology Software Tools, Inc. 2012). A keyboard was used for navigation through the experiment, as well as for entering responses to the verification question. Each item component (context, question, answer, and verification question) was presented on a separate screen. All items and instructions were in white print on a black backdrop except for the screen with the negative word, which had the text in green. Participants were instructed to say the word in

green aloud. They advanced through the experiment at their own pace, and the entire protocol took between 15 and 30 minutes.[12] Following the experiment, participants completed a brief (approximately five-minute) language history questionnaire online, via Google Docs.[13]

Oral responses were recorded using a Fostex DC-R302 recorder and a head-mounted Audix HT5 condenser microphone. The data were digitized at 44.1 kHz, 16 bit. Participants completed the experiment in a sound-attenuated booth, where they sat in front of a computer monitor by themselves.

2.4 Data Processing and Analysis

A total of 560 negative word tokens were recorded from the 20 participants whose productions we analyzed. However, the acoustic analysis was limited to items that received a correct answer to the verification question; i.e., those for which the oral answer was conceivably produced with the intended meaning according to the context provided. This resulted in a total of 504 tokens, which shows that participants tended to be accurate in their interpretation of the contexts. Another six tokens were discarded due to elision of a syllable or disfluencies, leaving a total of 498 tokens.

The data were segmented in Praat (Boersma & Weenink 2016). First, the beginning and end of the word were manually annotated via inspection of the synchronized waveform and

[12] E-Prime software collects reaction time information for self-paced tasks, but because items were not controlled for length, analysis of these data to draw inferences about processing difficulty would not be valid.

[13] The language history questionnaire was used to determine whether participants were native speakers of American English, and included questions on their location(s) of upbringing, family language background, knowledge of other languages, musical training, and history of language impairment. The results are reported above in section 2.1.

spectrogram. Then, syllable boundaries were marked following criteria in Turk et al. (2006). In all cases, intervocalic consonants were assumed to be onsets of the second syllable.

After segmentation, a series of measurements were extracted automatically. We extracted two relative measurements: relative duration of the stressed syllable (stressed syllable duration divided by the total word duration) and relative intensity of the stressed syllable (stressed syllable intensity divided by the total word intensity). The intensity parameters were set at: 100-Hz minimum pitch, 0-second time step, and subtract mean = yes. For more information, see the manual provided in Praat (Boersma & Weenink 2016).

These measurement values were z-score transformed by speaker. (Values were normalized separately.) Z-scored transformed measurements were analyzed in R (R Core Team 2015) using linear mixed effects regression (LMER) models (Bates et al. 2015).

In addition, time-normalized f0 values were extracted using ProsodyPro (Xu 2013), a Praat-based software developed to facilitate prosodic analysis of large corpora of speech data. Specifically, f0 values were extracted at ten equidistant points within each syllable, thus allowing us to compare different renditions of the same word across speakers and contexts. F0 values were also z-score transformed to allow us to compare across speakers. The resulting f0 curves were then analyzed statistically using smoothing spline (SS) ANOVAs (Gu 2014). SS ANOVA has been used in phonetic research to compare tongue shapes generated through ultrasound imaging (Davidson 2006), formant trajectories (Simonet et al. 2008; Nance 2014), as well as f0 curves (Mathes 2015). First, smoothing splines are fitted to each of the data sets being compared (here, NC vs. DN vs. Control). The smoothing splines plus the Bayesian confidence intervals are then plotted to visually compare the curves. The lack of overlap between the confidence intervals is interpreted as indicating a statistically significant difference between the curves (Davidson 2006).

Of the 504 target words produced, 168 target words (33% of the total) were annotated following the AmE_ToBI proposal (Beckman & Pierrehumbert 1986; Veilleux, Shattuck-Hufnagel & Brugos 2006). The tokens annotated were equally balanced across conditions (14 items × 4 negation words × 3 experimental conditions) and were randomly selected. The data were independently annotated by two ToBI trained transcribers who then consulted those cases that presented doubts. The results reported in section 2.5 are based on 163 tokens. Five tokens were discarded due to disfluencies.

The responses to the verification questions for the critical and control items (n = 28) were also analyzed statistically to determine whether participants were equally likely to give target responses in the DN, NC, and Control conditions. The data were analyzed in R using a general linear mixed effects regression model (GLMER; Bates et al. 2014).

The next section reports results of the acoustic analyses and ToBI annotation.

2.5 Results

2.5.1 ToBI Annotation

In total, we observed 13 different pitch accent and boundary tone configurations, yet, only one was found repeatedly across conditions. Table 1 shows the three most common pitch contours associated with each of the three conditions. The most common configuration was H* L-L%, characterized by a high f0 aligned with the stressed syllable and a fall in the posttonic syllable(s). This contour was found in 58% of the tokens annotated. The percentages by condition are shown in the table.

For the DN and NC conditions, the second and third most frequent contours are the same: H* L-H% and L+H* L-L%. The former presents a high f0 during the stressed syllable followed by a fall that does not reach such a low f0 as the H* L-L% but rather ends in a

sustained tone. The L+H* L-L% is also similar to the most frequent contour, but differs in the realization of the stressed syllable, which displays a rise instead of a flat H* tone.

In the Control condition, the second most common contour is the L* H-L%, which we also find with some frequency in the DN condition. The stressed syllable is produced with a low f0 that starts rising in the posttonic and stays flat in the following syllables. Finally, tokens in the Control condition were also realized with a L* L-L% contour: low f0 in the stressed syllable and in the posttonic syllable(s).

If we examine the pitch accents and boundary tones separately, it is worth noting that H* is the pitch accent most frequently found in our data (73.6% of the data), followed by L* (19.6%). Bitonal pitch accents are, thus, quite rare. The L-L% boundary tone is also the most representative of the data (69.9%).

Table 1. The three most common pitch contour configurations and their frequency (in counts and percentage) by condition.

Control		DN		NC	
H* L-L%	29 (52.7%)	H* L-L%	34 (61.8%)	H* L-L%	32 (60.4)
L* H-L%	6 (10.9%)	H* L-H%	5 (9.1%)	H* L-H%	8 (9.1%)
L* L-L%	5 (9.1%)	L+H*L-L% L* H-L%	4 (7.3%) 4 (7.3%)	L+H* L-L%	4 (7.5%)

2.5.2 Point Data

Figures 1 and 2 illustrate the normalized mean intensity and duration values for the negative words in the critical and control conditions.

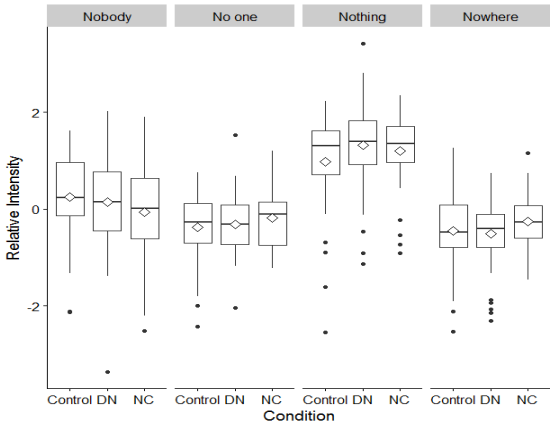

Figure 1. Relative intensity (z) of the stressed syllable in negative words by condition (Control, DN, NC)

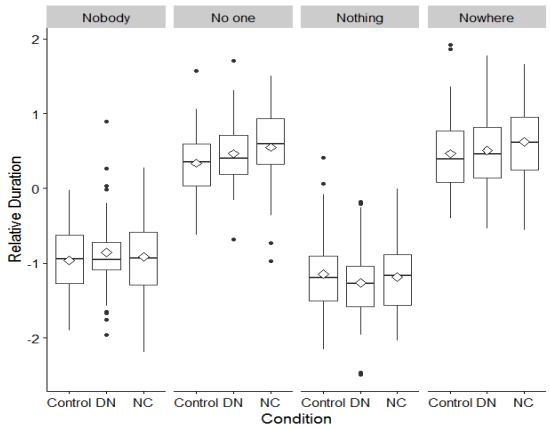

Figure 2. Relative duration (z) of the stressed syllable of negative words by condition (Control, DN, NC)[14]

[14] Fig. 2 shows relative duration of the stressed syllable with respect to the whole word. The differences that can be observed between the words are due to their segmental makeup (i.e., the total number of syllables the words have and the characteristics of the vowel nucleus of the stressed syllable). In *nothing*, the vowel is a monophthong ([ʌ]), so the duration of the syllable with respect to the word is expected to be shorter than for the other two disyllabic words that have a diphthong in the target syllable (*no one*, *nowhere*). *Nobody* also presents a shorter stressed syllable with respect to the word, but in this case it is due to the fact that the word has three syllables. Thus, differences across

The figures illustrate that, although there were differences between words, both relative intensity and relative duration of the stressed syllable with respect to the word were the same for each negative word across experimental conditions. The small variations that we observe seem negligible, and, in fact, statistical analyses confirm that. Two LMER models were fitted, one for relative stressed syllable duration and another one for relative stressed syllable intensity. Both models had Condition as a fixed effect and random intercepts and slopes for Participant and Item. We found no significant difference in duration between the Control Condition and DN (β = .05, SE = .22, p(z) = .21) or NC (β = .03, SE = .22, p(z) = .14), and no significant difference in intensity between the Control Condition and DN (β = .06, SE = .23, p(z) = .24) or NC (β = .12, SE = .23, p(z) = .5).

2.5.3 F0 Curves

As regards the analysis of f0 curves, Figures 3 and 4 show the smoothing splines plus the Bayesian confidence intervals for the target negation words. Since *nobody* was the only three syllable word, it was analyzed separately given that the relevant syllables (stressed and final) would not line up if all the data were analyzed together.

words do not indicate that the *no-* in *nobody* is shorter in absolute terms than the *no-* in *no one* and *nowhere*, only in relative terms because the word is longer.

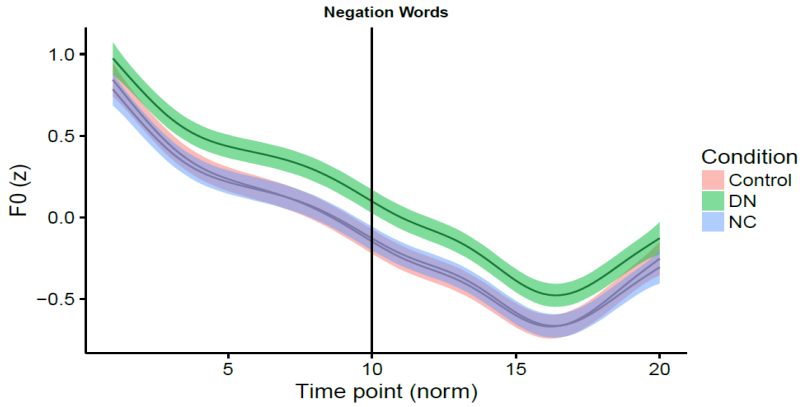

Figure 3. Smoothing splines and 95 % Bayesian confidence intervals for the f0 curves corresponding to two-syllable negative words *no one*, *nothing*, and *nowhere* by condition. The vertical line indicates the syllable boundary.

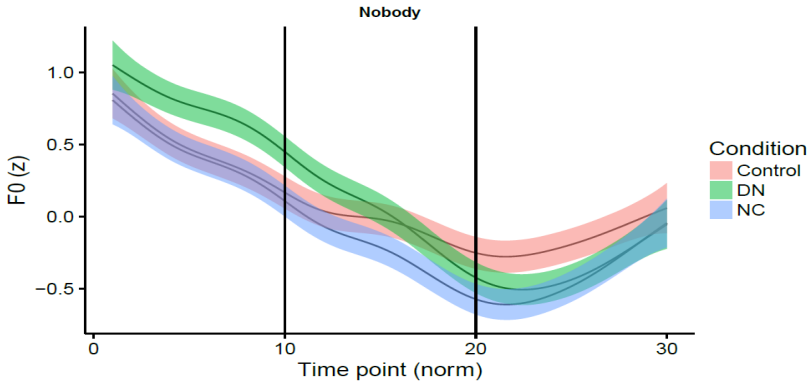

Figure 4. Smoothing splines and 95 % Bayesian confidence intervals for the f0 curves corresponding to the three-syllable negative word *nobody* by condition. Vertical lines indicate syllable boundaries.

As shown in Figure 3, the stressed (first) syllable of disyllabic negation words shows a falling f0 movement that continues to fall during the posttonic until it reaches its minimum around the final syllable mid-point, where it increases slightly. We observe almost complete overlap between the Control and NC conditions. As for the DN condition, while it presents the same overall contour shape, the normalized f0 curve does not overlap with the other two, given that it is significantly higher.

Figure 4 shows that *nobody*, with three syllables, presents the same nuclear contour as the two syllable words (i.e., the overall shape of the f0 curve is the same), with a progressive f0 fall from the beginning of the word until it reaches the f0 minimum within the last syllable, at which point there is a change in f0 direction. Like in Figure 3, the DN condition exhibits higher f0 in the first syllable (the stressed syllable, site of the nuclear pitch accent) than the other two conditions. In addition, it also presents a higher f0 than the NC condition in the second syllable, although not in the third syllable. In that last syllable, the Control condition fails to reach the minimum f0 reached in the other two conditions. The f0 rise in the second part of the last syllable is much less steep in the Control condition than in the other two conditions.

2.6 Verification Question Response Data

Figure 5 illustrates participant responses to the true or false verification questions, included after each experimental item to evaluate whether participants interpreted the contexts as intended. Percentages of expected responses for the single negative Control, DN, and NC condition are shown.

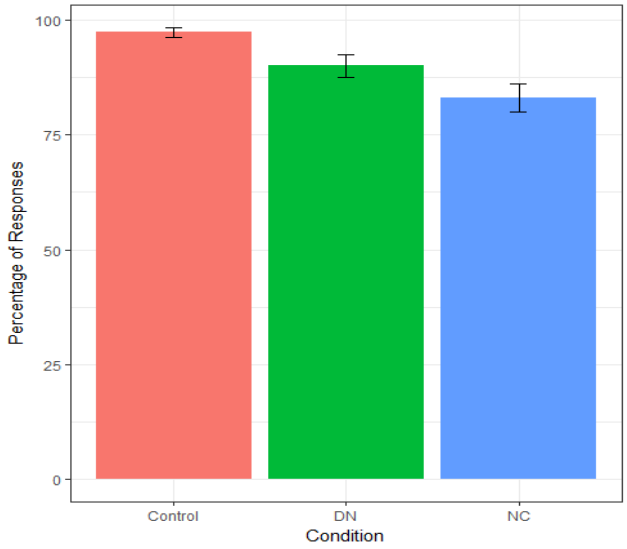

Figure 5. Percentage of target-like responses to Control, DN, and NC verification questions.

Figure 5 shows that responses were on target most of the time in all three negation conditions. Performance was nearly at ceiling in the single negative Control condition, in which responses were target-like 97 % of the time. In the DN condition responses were target-like 90 % of the time, and for NC participants gave 83 % target-like responses. A GLMER model comparing response rates for NC and DN item pairs (e.g. (8) vs. (9)) with random intercepts for Participant and Item was performed to determine the effect of condition on target response.[15] This revealed that DN and NC target response rates were not equivalent, and that participants were significantly more likely to give non-target like answers with the NC items than they were with DN (β = −.66, SE = .30, p(z) < .05). This result suggests that the True/False verification questions following the critical items were easier to answer in DN contexts than they were in contexts that elicited a single negation reading.

Because they have different contexts, the control items could not be paired with the critical items. Therefore, a separate GLMER was run to compare the critical conditions against the single negative controls. This revealed that participants gave significantly fewer target-like responses in both the DN (β = −1.39, SE = .51, p(z) < .01) and the NC (β = −2.01, SE = .49, p(z) < .001) conditions than in the Control condition.

Summarizing, our acoustic results show that NC and single negation interpretations of negative indefinites have similar intonational form and identical f0, while DN, though similar in intonational form to NC and single negation, is marked by higher fundamental frequency. In the comprehension task, participants performed well on all conditions, but had greater

[15] The sparsity of data prevented us from using a maximal random effects structure (Barr et al. 2013). Additionally, a log likelihood ratio test confirmed that adding the interaction between condition and negative word does not significantly change the model (χ^2 (3) = 2.77, p > .05). We therefore calculated only the effect of Condition on correctness, taking Participant and Item as random variables.

difficulty with DN than with the single negative control items, and had the lowest frequency of target-like responses on the NC condition.

3. Naturalness Ratings

To better understand the results of the production task, we conducted a second experiment, in which we asked a different but demographically similar group of participants to provide naturalness ratings of our stimuli. Because it did not involve collecting acoustic data, this experiment was conducted online. In this section, we describe the methodology and present the results of the naturalness ratings experiment. The results of both the production and the naturalness ratings experiments are discussed subsequently, in Section 4.

3.1 Participants

47 undergraduates were recruited from two introductory linguistics classes at a public university in the United States, and they received extra credit toward their course for completing the survey. We discarded data from two non-native speakers, leaving us with data from 45 American English speakers ranging between 18 and 28 in age (M = 19.47, SD = 1.85). Following the experiment, participants completed a language history questionnaire. This revealed that participants came from eight different U.S. states, including California (1), Connecticut (2), Maryland (4), Massachusetts (1), New Jersey (11), New York (5), Pennsylvania (19), and South Carolina (2). 33 participants reported that they grew up in suburban areas, seven reported having rural upbringings, and five were from urban areas. 26 reported to having at least some knowledge of another language, with Spanish (10 participants), French (5), German (4), and Italian (4), being the most common.

3.2 Materials and Procedure

The items used in the naturalness ratings study were the same as those employed in the production study. The survey was programmed and distributed using Qualtrics (2017). Context, question, and target word were presented first, and participants were asked to rate the naturalness of the target word on a scale of 1–5. A separate screen with the True/False verification question immediately followed each item.

3.3 Results

3.3.1 Verification Question Results

Figure 6 illustrates participant responses to the true or false verification questions in the naturalness ratings experiment.

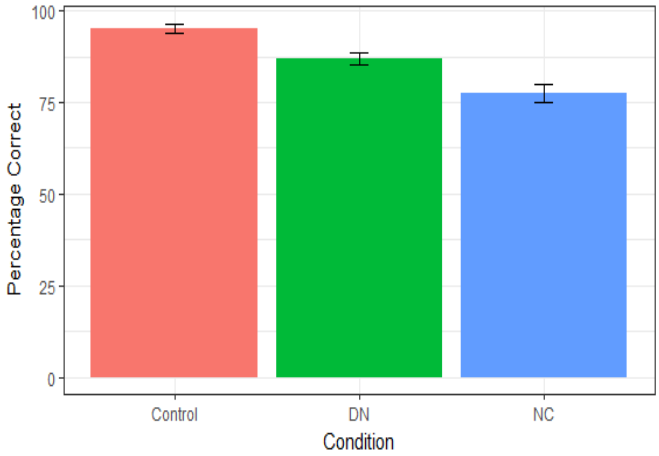

Figure 6. Percentage of correct responses to verification questions in the naturalness task by condition.

Figure 6 shows a pattern similar to that of the production task. As with the production task, responses were on target most of the time in all three conditions. In this case, means were slightly lower overall, at 95% for the Control condition, 86% for DN, and 78% for NC. A GLMER model with random intercepts for Participant and Item was used to compare response rates for NC and DN item pairs. The model revealed that, like in the production task, participants were significantly more likely to give non-target like answers with the NC items

than they were with DN (β = −.99, SE = .22, p(z) < .001). A GLMER comparing the critical conditions against the controls revealed that participants gave significantly fewer target-like responses in both the DN (β = −1.17, SE = .6, p(z) < .001) and the NC (β = −2.18, SE = .59, p(z) < .001) conditions than in the Control condition.

3.3.2 Naturalness Ratings

Figure 7 shows the naturalness ratings associated with items that obtained correct responses to the verification questions.

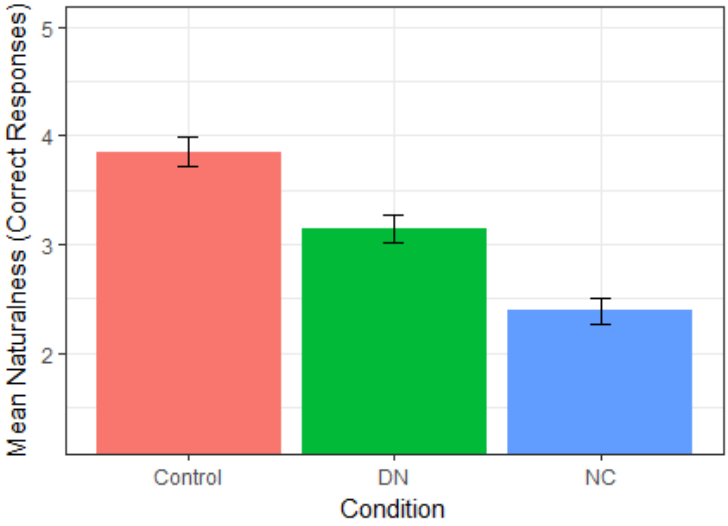

Figure 7. Naturalness ratings (on a scale of 1–5) for items with correct answer to verification question by condition.

Overall, participants' mean naturalness ratings on the 1–5 Likert scale were 3.85 for the Controls, 3.13 for DN, and 2.39 for NC. A linear mixed effects model comparing naturalness ratings for items with target-like responses in the critical conditions, with random slopes for Participant and Item, revealed that participants gave significantly lower ratings for NC than for DN (β = −1.49, SE = .09). A comparison of the critical conditions with the Control condition

revealed that both DN (β = −1.21, SE = .11) and NC (β = −2.70, SE = .11) were rated significantly lower than the controls.

4. Discussion

4.1 Prosody and the Meanings of Negative Indefinites

The results of our acoustic analysis of the data from the production experiment demonstrate that, like in prototypical NC languages, denial is the marked interpretation for single negative indefinites in question-answer pairs in English, as compared with the single negative meaning of both the single negation and the Control items. The markedness of DN was encoded by fundamental frequency alone. The overall shape of the prosodic contour for DN, NC, and the single negative controls was found to be the same, but f0 was significantly higher for the DN items than for both the NC and control items, which had overlapping f0 values. The prosodic transcription of the data confirms that there is a strong tendency for tokens in all three conditions to be realized with the same intonational contour (H* L-L%), although it is true that a variety of contours were found and that some of them are more frequent in one condition than in the others. Still, the H* L-L% contour was found in over 50% of the tokens in all three conditions.

These results are both similar to and different from the results reported for Catalan in Espinal & Prieto (2011), and for Spanish and Catalan in Prieto et al. (2013) and Espinal et al. (2016), all of which used a perception task to examine the behavior of single negative words in question-answer pairs. Our results pattern with these previous studies' results in that a more marked prosody is used in English to convey a denial interpretation rather than a single negation one, but, unlike in Spanish and Catalan, the difference does not seem to be phonological, but rather, phonetic. That is, English does not seem to employ a different combination of pitch accent and boundary tone, but rather f0 register. As mentioned, the most

frequent contour across conditions was H* L-L%. Our data did not reveal the use in English of different tonal configurations to signal the different meanings analyzed.

The crucial point for our purposes is that, in Spanish and Catalan, it is the DN reading that is associated with a more marked prosody (with the complex boundary tone vs. the more simple boundary tone of the NC reading), and not NC. Taking all the data production data together, in our case, the more marked reading (DN) is rendered with an overall higher f0 register. The acoustic properties of this markedness are at least partially consistent with the results in Breen et al. (2010), who found that English contrastive focus is marked by a higher f0, greater intensity, and longer duration than regular, non-contrastive focus. English thus behaves like prototypical NC languages in marking DN prosodically, with expected crosslinguistic differences in terms of how this marking is realized (Ladd 2008).

Our acoustic results are also informative with respect to the broadly assumed English-internal division between NC varieties and Standard English, thought to be prototypically DN. Our participants reliably distinguished prosodically between single negation and DN readings of negative indefinites, aligning the single negation responses to negative questions with their equivalent responses to non-negative questions. This suggests that like Spanish and Catalan, Standard English is a language that generates both single and double negation readings of negative words, each with its own set of prosodic, syntactic, and pragmatic conditions.

In their verification question responses, which immediately followed their productions, although participants displayed consistently high rates of target responses (97 % for the Controls, 90 % for DN, and 83 % for NC), there was a significant difference in their responses across all three conditions. This result shows that the negation in the question played a role in the interpretation of the fragment answer. We acknowledge that, because the single word response by definition has no immediate surrounding context, a correct response to the verification question relied crucially on participants' comprehension of the preceding context.

Because participants answered the verification questions correctly the overwhelming majority of the time in all three conditions, and because those correct responses displayed a reliable marking of DN relative to single negation in both the NC and the Control conditions, we are led to conclude on the basis of our data that Standard English speakers access and produce both single and double negation interpretations of negative word responses to negative questions.

Negation is known to be difficult to process, as determined by both online and offline psycholinguistic measures (see Tian & Breheny 2016 for a review). The fact that our participants gave significantly fewer target-like responses in the DN condition than for the Controls provides an offline measure which suggests that denial negation, a form of metalinguistic negation (at least in the context of our study), is more difficult to comprehend in context than single, or descriptive negation. This offline result corroborates the online eye-tracking measures in Orenes et al. (2016), which show that participants are slower to process metalinguistic negation than simple negation. This is in contrast with the eye-tracking results in Noh et al. (2013), in which descriptive and metalinguistic negation were found to induce equivalent processing loads. Our results therefore contribute additional empirical information to inform this debate, suggesting that metalinguistic negation, when presented as denial in the context of a negative response to a negative question, may in fact be more difficult to process than single negation that does not interact with a negation in a previous utterance.

With respect to performance on the NC condition, our analysis raises the question of why, if participants can access both single negation and DN interpretations of negative responses to negative questions, the verification question responses for the NC items were significantly less target-like than both the DN items and the controls. The production experiment design and results point to two possible explanations for this. One possibility is that participants' degraded performance on the NC condition is an effect of the prescriptive ban on NC in English. While the acoustic data reflect participants' online expression of the NC items

as single negation, the offline comprehension measure allows for time to reflect on the acceptability of using a negative response to a negative question in the formal context of a university laboratory. It is therefore possible that, given time to consider their response, degraded performance on the NC condition is a reflex of normative pressures, and not a true grammatical effect.

Another possible explanation is that the single negation response interacts with the presuppositions projected by the question in a manner that yields some level of pragmatic infelicity. Consider again the NC item in (8) above, repeated here as (11):

(11) Context: You and your roommate pay different bills each month. This month you have too little money to pay bills.
Question: Your roommate asks: What didn't you pay?
Response: Nothing.

In Section 1.3 we discussed the intuition that the negative question may be associated with presuppositions which are "soft" or defeasible in the sense of Abusch (2010). In this case, the question *what didn't you pay?* seems to presuppose that (i) there exist some bills that were paid, and (ii) there exist some bills that were not paid. To explain the why participants were worse at answering the verification question with NC than they were with DN, we suggest that the negation introduced by the question, which scopes below the existence presupposition included in (ii), makes the presupposition more salient and therefore less defeasible. We further hypothesize that this effect yields a level of pragmatic infelicity that impacts interpretation. This latter explanation appears to be better supported by the naturalness ratings we reported in section 3, which we turn to directly.

4.2 Naturalness ratings, single negation, and DN

As reported in Section 3, in an experiment separate from the production experiment, but conducted with a very similar population, naturalness ratings of our stimuli on a 1–5 scale were found to be higher for the Control condition than for both NC and DN. Furthermore, as regards the two critical conditions, NC items were perceived to be less natural on average than DN items. If nothing else is said, this result seems to support both possible interpretations of our verification question results from the production experiment. Participants degraded ratings for the NC condition could be lower because they display the effects of prescriptive judgments, and they could also be lower because participants found them to be less felicitous due to interference from the question negation with the defeasibility of the "soft" existence presupposition, which is required for the single negation interpretation.[16]

Although our naturalness ratings do not provide clear evidence to distinguish between these two possibilities, we argue on theoretical grounds and on the basis of independent evidence that the infelicity explanation is superior to the prescriptivist account. To understand, we must first recall the difference between ellipsis and non-ellipsis analyses of negative fragment answers to negative questions, discussed in more detail in section 1.2. Recall that under ellipsis accounts, the negative fragment answer is the overt realization of what is underlyingly a full clausal structure (i.e., *I didn't pay nothing*). This means that for the NC response, an NC relation is established underlyingly between the two negations within the clause. In Espinal and Tubau's (2016) alternative account, which does not involve ellipsis, there is no such relation established within the clause, because there is no second negation (and in fact, there is no clause). Instead, the negation in the answer "combines" with the negation in

[16] Recall also that participants' performance on the verification questions, though still very good, was slightly worse following the naturalness ratings than it was following the production task, suggesting that the act of judging the sentence somehow made it more difficult to interpret.

the question, establishing a sort of cross-utterance concord relation. Consider now the reasoning that participants must undertake to formulate their prescriptive judgment. On the ellipsis analysis, they must make the conscious realization that, though absent from view, there is another negation present in the answer. Subsequently, they must deem that invisible negation to be violating the prescriptive rule that in English, two negatives must equal a positive (Lowth 1762), and therefore judge the sentence to be unacceptable, and significantly less so than its DN counterpart, in which the "invisible" negation does in fact contribute to interpretation. On the non-ellipsis account, the prescriptivist account suggests that participants form an analogy in which the combination of two negations across utterances has the same social status as NC within a sentence, and thus judge it unacceptable. Both of these strategies seem highly unparsimonious, at best, and the analogy required by the non-ellipsis account seems to make the prescriptivist explanation even less likely.

Recall now our section 1.1 discussion of Blanchette (2017), in which Standard English speakers were shown to reliably rate constructions with the marker *–n't* and a negative object significantly higher in NC contexts than in DN contexts. Note that, under an ellipsis account, Blanchette's results directly contradict the results of our naturalness ratings study. This is because all of our stimuli included contexts in which, under an ellipsis account, the negative phrase is underlyingly a negative object (e.g., *I didn't pay nothing*). On the basis of Blanchette's results, an ellipsis account therefore predicts that participants will prefer NC over DN contexts, and yet our participants' naturalness ratings displayed the reverse pattern. We therefore suggest that our naturalness ratings data provide indirect support for Espinal & Tubau's (2016) account of negative fragment answers to negative questions as single phrases with no underlyingly elided structure. We further suggest that in the context of previous work on English NC and DN, adopting a non-ellipsis account allows us to conclude from our data that negative questions compromise the conditions for soft presupposition cancellability, making single negative

fragment answers in English less "natural" or felicitous than they are in non-negative questions, and less natural in negative questions than their DN counterparts.

5. Conclusion

On the basis of the acoustic and behavioral data presented in this paper, we have argued that Standard English is a language that generates both single negation and DN interpretations of negative indefinite fragment answers to negative questions, with DN as the prosodically marked form. The results of our acoustic analysis, coupled with the results of our prosodic transcription, clearly show that DN, which serves as a denial negation in our stimuli, is marked by a higher fundamental frequency than its single negative counterparts, even if it is frequently produced with the same intonational contour that we find in the other two conditions. Results of both the verification questions and the naturalness rating task indicate that there is an asymmetry between the two experimental conditions, NC and DN, such that DN items receive a higher number of correct responses and are perceived to be more natural. We attribute the degraded acceptability of NC to an effect of the negation introduced in the question, which serves to reduce the defeasibility of a presupposition introduced by the question (in the sense of Abusch 2010), which makes the single negation interpretation more difficult to access.

Acknowledgments

Thanks to Clara Cohen and Matthew Carlson for help planning the statistical analyses, and to Meagan Lynn Daly for assisting with data collection and processing. Thanks also to Chris Collins for comments on an earlier version of this manuscript, and to two anonymous reviewers for helpful feedback and comments.

References

Abusch, Dorit. 2010. Presupposition triggering from alternatives. *Journal of Semantics* 27: 37–80.

Alonso-Ovalle, Luis & Elena Guerzoni. 2004. Double Negation, Negative Concord and Metalinguistic Negation. In Mary Andronis, Erin Debenport, Anne Pycha and Keiko Yoshimura (eds.), *Proceedings of the Chicago Linguistics Society Meeting, 38–1: The Main Session*, 15–31. Chicago, IL: CLS Publications.

Anderwald, Liesolette. 2002. *Negation in Non-Standard British English: Gaps, Regularizations, and Asymmetries*. London: Routledge.

Anderwald, Liesolette. 2005. Negative Concord in British English dialects. In Yoko Iyeiri (ed.), *Aspects of English Negation*, 113–137. Amsterdam: John Benjamins.

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

Bates, Douglas, Martin Maechler, Ben Bolker, and Steve Walker. 2014. Lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–10. URL: https://github.com/lme4/lme4pureR.

Beckman, Mary E. & Janet Pierrehumbert. 1986. Intonational structure in English and Japanese. *Phonology Yearbook* 3: 255–310.

Belnap, Nuel D. Jr. 1966. Questions, answers, and presuppositions. *Journal of Philosophy* 63: 609–611.

Biberauer, Theresa & Hedde Zeijlstra. 2012. Negative Concord in Afrikaans: Filling a typological gap. *Journal of Semantics* 29: 345–271.

Blanchette, Frances. 2013. Negative concord in English. *Linguistic Variation* 43: 1–47. DOI: http://dx.doi.org/10.1075/lv.13.1.01bla.

Blanchette, Frances. 2015. English Negative Concord, Negative Polarity, and Double Negation. Doctoral Dissertation, CUNY Graduate Center.

Blanchette, Frances. 2017. Microsyntactic variation in English Negative Concord. *Glossa: A Journal of General Linguistics* 2: 1–32.

Boersma, Paul & David Weenink. 2016. Praat: Doing phonetics by computer. URL: http://www.praat.org/.

Breen, Mara, Evelina Fedorenko, Michael Wagner, & Edward Gibson. 2010. Acoustic correlates of information structure. *Language and Cognitive Processes* 25: 1044–1098.

Collins, Chris, & Paul M. Postal. 2014. *Classical NEG Raising: An essay on the syntax of negation*. Cambridge, Massachusetts: MIT Press.

Cole, Jennifer. 2015. Prosody in context: A review. *Language, Cognition and Neuroscience* 30: 1–31. DOI: 10.1080/23273798.2014.963130.

Dayal, Veneeta. 2016. *Questions*. Oxford: Oxford University Press.

Déprez, Viviane. 2000. Parallel (a)symmetries and the internal structure of negative expressions. *Natural Language and Linguistic Theory* 18: 253–342.

Déprez, Viviane. 2011. Atoms of negation: An outside-in micro-parametric approach to Negative Concord. In Richard Ingham & Pierre Larrivée (eds.), *The Evolution of Negation: Beyond the Jespersen Cycle*,221–272. Berlin: Mouton de Gruyter.

Déprez, Viviane, Susagna Tubau, Anne Cheylus, & M. Teresa Espinal. 2015. Double negation in a negative concord language: An experimental investigation. *Lingua* 163: 75–107.

Déprez, Viviane & Jeremy Yeaton (to appear). French Negative Concord and Discord: An experimental investigation of contextual and prosodic disambiguation. *Proceedings of LSRL 46*.

De Swart, Henriette. 2010. *Expression and Interpretation of Negation. Studies in Natural Language and Linguistic Theory, Vol 77*. Dordrecht: Springer. DOI: 10.1007/978-90-481-3162-4_4.

De Swart, Henriette & Ivan Sag. 2002. Negation and Negative Concord in Romance. *Linguistics and Philosophy* 25: 373–417.

Davidson, Lisa. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *The Journal of the Acoustical Society of America* 120(1): 407–415. DOI: http://dx.doi.org/10.1121/1.2205133.

Dunstan, Stephany Brett & Audrey J. Jaeger. 2015. Dialect and influences on the academic experiences of college students. *The Journal of Higher Education* 86: 777–803.

Espinal, M. Teresa & Pilar Prieto. 2011. Intonational encoding of Double Negation in Catalan. *Journal of Pragmatics* 43: 2392–2410.

Espinal, M. Teresa & Susagna Tubau. 2016. Interpreting argumental n-words as answers to negative questions. *Lingua* 177: 41–59.

Espinal, M. Teresa, Susagna Tubau, Joan Borrás-Comes, & Pilar Prieto. 2016. Double Negation in Catalan and Spanish: Interaction between syntax and prosody. In Pierre Larrivée & Chungmin Lee (eds.), *Negation and Polarity: Experimental Perspectives*, 145–176. Dordrecht: Springer.

Geurts, Bart. 1998. The mechanisms of denial. *Language* 74: 274–307.

Giannakidou, Anastasia. 1998. *Polarity Sensitivity as (Non)veridicality*. Amsterdam: John Benjamins.

Goodhue, Daniel & Michael Wagner. 2018. Intonation, *yes* and *no*. *Glossa: a journal of general linguistics* 3:1–45.

Green, Lisa. 2002. *African American English: A Linguistic Introduction*. Cambridge: Cambridge University Press.

Gu, Chong. 2014. Smoothing spline ANOVA models: R Package GSS. *Journal of Statistical Software* 58(1): 1—25. DOI: 10.18637/jss.v058.i05.

Hamblin, Charles L. 1973. Questions in Montague English. *Foundations of Language* 10: 41–53.

Harrison, Simon. 2009. Grammar, Gesture, and Cognition: The Case of Negation in English. Doctoral Dissertation, Université de Bordeaux 3, Bordeaux.

Henry, Alison. 2016. Acquiring language from variable input. *Linguistic Variation* 16: 131–150.

Holmberg, Anders. 2016. *The syntax of yes and no*. Oxford: Oxford University Press.

Horn, Laurence R. 1989 [2001]. *A Natural History of Negation*. Chicago, IL: University of Chicago Press.

Horn, Laurence. 2010. Multiple negation in English and other languages. In Laurence Horn (ed.), *The Expression of Cognitive Categories: Expression of Negation*, 117–148. Berlin: Walter de Gruyter.

Horton, Ho'omana Nathan. 2017. Language discrimination on campus: Ratings of and attitudes toward student writing with African American English. *Proceedings of the Linguistic Society of America* 2: 1–11. DOI: http://dx.doi.org/10.3765/plsa.v2i0.4041.

Johnson, David & Lewis VanBrackle. 2011. Linguistic discrimination in writing assessment: How raters react to African American "errors," ESL errors, and standard English errors on a state-mandated writing exam. *Assessing Writing* 17: 35–54.

Kallel, Amel. 2007. The loss of Negative Concord in Standard English: Internal factors. *Language Variation and Change* 19: 27–49.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1: 3–44.

Katz, Jerrold J. & Paul M. Postal. *An Integrated Theory of Linguistic Descriptions*. Cambridge, MA: MIT Press.

Kaufman, Anita. 2002. Negation and prosody in British English: a study based on the London-Lund Corpus. *Journal of Pragmatics* 34: 1473–1494.

Kramer, Ruth & Kyle Rawlins. 2009. Polarity particles: An ellipsis account. In *The Proceedings of the Northeast Linguistics Society (NELS) 39*, 479–472.

Krifka, Manfred. 2013. Response particles as propositional anaphors. In The Proceedings of Semantics and Linguistic Theory (SALT) 23: 1–18. DOI: https://doi.org/10.3765/salt.v23i0.2676.

Ladd, Robert. 2008. *Intonational Phonology*, 2nd Edition. Cambridge: Cambridge University Press.

Ladusaw, William A. 1992. Expressing negation. In Chris Barker & David Dowty (eds.), *Proceedings of SALT II, Ohio State Working Papers in Linguistics 40*, 237–259. Columbus, Ohio: Ohio State University.

Lawrence, Michael. 2013. ez: Easy analysis and visualization of factorial experiments. R package version 4.2-0.

Lewis, Shevaun & Colin Phillips. 2015. Aligning Grammatical Theories and Language Processing Models. *Journal of Psycholinguist Research* 44: 27-46. DOI 10.1007/s10936-014-9329

Liberman, Mark & Ivan Sag. Prosodic form and discourse function. In Michael W. La Galy, Robert A. Fox, & Anthony Bruck (eds.), *Papers from the Tenth Regional Meeting of the Chicago Linguistic Society*, 416–427. Chicago: Chicago Linguistic Society.

Longobardi, Giuseppe. 2014. Theory and experiment in parametric minimalism: The case of Romance negation. In Rob Pensalfini, Myfany Turpin, & Diana Guillemin

(eds.), *Language Description Informed by Theory*, 217–262. Amsterdam: John Benjamins.

Lowth, (Bishop) Robert. 1762. *A Short Introduction to English Grammar*. London: J. Hughs.

Mathes, Timothy K. 2015. Consonant-Tone Interaction in the Khoisan Language Tsua. Doctoral Dissertation, New York University. https://ling.auf.net/lingbuzz/002725.

Merchant, Jason. 2001. *The Syntax of Silence: Sluicing, Islands and the Theory of Ellipsis*. Oxford: Oxford University Press.

Nance, Claire. 2014. Phonetic variation in Scottish Gaelic laterals. *Journal of Phonetics* 47: 1–17. DOI: http://dx.doi.org/10.1016/j.wocn.2014.07.005.

Nevalainen, Terttu. 2006. Negative Concord as an English "Vernacular Universal": Social history and linguistic typology. *Journal of English Linguistics* 34: 257–278.

Noh, Eun-Ju, Hyeree Choo & Sungryong Koh. 2013. Processing metalinguistic negation: Evidence from eye-tracking experiments. *Journal of Pragmatics* 57: 1–18.

Orenes, Isabel, Linda Moxey, Christoph Scheepers, & Carlos Santamaría. 2016. Negation in context: Evidence from the visual world paradigm. *The Quarterly Journal of Experimental Psychology* 69: 1082–1092.

Pierrehumbert, Janet B. 1980. The Phonology and Phonetics of English Intonation. Doctoral Dissertation, Massachusetts Institute of Technology.

Prieto, Pilar, Joan Borràs-Comes, Susagna Tubau & M. Teresa Espinal. 2013. Prosody and gesture constrain the interpretation of double negation. *Lingua* 131: 136–150. DOI: http://dx.doi.org/10.1016/j.lingua.2013.02.008.

Puskás, Genoveva. 2012. Licensing Double Negation in NC and non-NC languages. *Natural Language & Linguistic Theory* 30: 611–649.

Psychology Software Tools, Inc. 2012. E-Prime 2.0. URL: http://www.pstnet.com.

Qualtrics. 2017. Version January 2018. Provo, Utah. URL: http://www.qualtrics.com.

R Core Team. 2015. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org.

Roloefson, Floris & Donka Farkas. 2015. Polarity particle responses as window onto the interpretation of questions and assertions. *Language* 91:359–514.

Simonet, Miquel, Marcos Rohena-Madrazo, & Mercedes Paz. 2008. Preliminary evidence for incomplete neutralization of coda liquids in Puerto Rican Spanish. In: Colantoni, Laura and Jeffrey Steele (Ed.), *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, 72—86. Somerville, MA: Cascadilla Proceedings Project. URL: www.lingref.com, document #1715.

Smith, Jennifer. 2001. Negative concord in the Old and New World: Evidence from Scotland. *Language Variation and Change* 13: 109–134.

Temmerman, Tanja. 2012. Multidominance, Ellipsis, and Quantifier Scope. Doctoral Dissertation, Utrecht University.

Thornton, Rosalind, Anna Notley, Vincenzo Moscati, & Stephen Crain. 2016. Two negations for the price of one. *Glossa: A Journal of General Linguistics* 45: 1–30.

Tian, Ye & Richard Breheny. 2016. Dynamic pragmatic view of negation processing. In. Pierre Larrivée & Chungmin Lee (eds.), *Negation and Polarity: Experimental Perspectives*, 21–43. Switzerland: Springer International Publishing.

Tubau, Susagna. 2016. Lexical variation and Negative Concord in traditional dialects of British English. *The Journal of Comparative Germanic Linguistics* 19: 143–177.

Turk, Alice, Satsuki Nakai, & Mariko Sugahara. 2006. Acoustic segment durations in prosodic research: A practical guide. In Stefan Sudhoff, Denisa Lenertová, Roland Meyer, Sandra Pappert, Petra Augurzky, Ina Mleinek, Nicole Richter, & Johannes Schließer (eds.), *Methods in Empirical Prosody Research*, 1–27. Berlin: Walter de Gruyter.

Veilleux, Nanette, Stefanie Shattuck-Hufnagel, & Alejna Brugos. 2006. Transcribing prosodic structure of spoken utterances with ToBI. Massachusetts Institute of Technology online course. URL: https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-911-transcribing-prosodic-structure-of-spoken-utterances-with-tobi-january-iap-2006/index.htm.

Wallage, Phillip. 2012. Negative inversion, Negative Concord and sentential negation in the history of English. *English Language and Linguistics* 16.1: 3–33.

Watanabe, Akira. 2004. The genesis of Negative Concord: Syntax and morphology of negative doubling. *Linguistic Inquiry* 35: 59–612.

Wickham, Hadley. 2009. *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer. URL: http://had.co.nz/ggplot2/book.

Wolfram, Walt & Donna Christian. 1976. *Appalachian Speech*. Arlington, VA: Center for Applied Linguistics.

Xu, Yi. 2013. ProsodyPro—A tool for large-scale systematic prosody analysis. In *Proceedings of Tools and Resources for the Analysis of Speech Prosody*: 7–10. Aix-en-Provence, France.

Yaeger-Dror, Malcah. 1995. Intonational prominence on negatives in English. *Language and Speech* 28: 197–230.

Yaeger-Dror, Malcah. 1997. Contraction of negatives as evidence of variance in register-specific interactive rules. *Language Variation and Change* 9: 1–36.

Zanuttini, Raffaella. 1991. Syntactic properties of sentential negation: A comparative study of Romance languages. Doctoral Dissertation, University of Pennsylvania.

Zeijlstra, Hedde. 2004. Sentential Negation and Negative Concord. Doctoral Dissertation, University of Amsterdam.