

Non-local Contexts Help Resolve Ambiguity

Daniel Krugman and Carl Vogel
Computational Linguistics Group
Trinity College Dublin Dublin 2, Ireland
d.delaney.krugman@gmail.com & vogel@tcd.ie

1 Background

This paper addresses nonlocal context effects in the interpretation of ambiguous utterances in natural language. We examine *equivocation* as a form of discourse ambiguity and demonstrate that nonlocal contexts can resolve ambiguity by providing a method for exploring the effects of global context. Of particular relevance is that the locus of ambiguity within the texts analyzed is within and across quotations included in larger texts that are representative of summaries of speeches as reported in newspapers. This research has relevance to sentiment analysis through the ramifications that sentiment relevant to financial markets cannot necessarily be detected from quoted texts alone, even when the text quoted in the article is that of the Federal Reserve Board chair. We think it safe to say that most research on sentiment analysis does not distinguish between direct text and text present indirectly via quotation. The texts we use as experimental items in our study involve a mixture of quoted and nonquoted statements of Alan Greenspan, texts which are relevant to domain-specific decision making. The results we report suggest that sentiment analysis research is mistaken if it does not parse for quotational contexts of sentiment bearing words. Our results show that nonlocal contexts strongly influence decision making behavior in response to ambiguous texts.

One thread of psycholinguistic research on ambiguity resolution is in the effect of context. In many cases, the processing oriented studies are focused on rapid real-time presentation of ambiguous statements, and corresponding probes. In such research, context is generally understood in very local terms. Neighboring words or a window of sentences on either side of an experimental item. Certainly, it is a common assumption that a window of a few sentences on either side assuredly resolves contexts for most natural interpreters. In this paper we report on an experiment in human response to ambiguous language in a realistically contextualizing situations. We find that sentential context can be over-ridden by larger scenarios. The research is motivated by the independent ques-

tion of ambiguity in the speeches of public servants, like the Chairman of the Federal Reserve Board.

In the remainder of this background section characterize psycholinguistic research that attends to context. We then provide a subjective analysis of ambiguity in some of the texts Greenspan has produced. We present an empirical study using a design based on a rich notion of context, manipulating context to identify influences of material beyond the discourse level on interpretation of the same ambiguous texts. We judge the influence on the basis of participant decisions to ‘buy’ or ‘sell’ when interpreting the speeches in corresponding contexts. In fact, it is possible to assess the speeches as patently ambiguous in that opposing decisions can be induced with manipulated contexts outside the texts of the speeches themselves.

Context Effects in Psycholinguistics

Altmann, Garnham, and Henstra (1994) found that syntactic disambiguation was sensitive to subtle changes in the semantic context. Sentences focused on object- versus subject-relative clauses such as 1 and 2 respectively. They found that reading times did differ between the two in that the object-relative sentences were processed faster than the subject-relative ones in a null context. However, these differences disappeared if the sentences were placed in contexts that supported the subject-relative interpretation (Altmann et al., 1994).

- (1) The politician told the woman that he had been meeting that he was going to see the minister.
- (2) The politician told the woman that had been meeting him that he was going to see the minister.

For instance, if the context for (2) went as follows, *A politician was talking to two women. He was telling them about his concern for the environment. One of the women had been meeting him regularly. The other woman hadn't met him before*, there would be no difference in reading times between the two sentences.

Altmann et al. (1994) suggested that, though this was not direct evidence, it did support referential hypothesis in that the number of possible referents present in the context can play a role in the relative speed of processing. While there may be a preference toward the simpler construction in a null context, an appropriate context can eliminate any difference in processing time. This lack of any difference again suggests that semantics and syntax are able to interact during interpretation.

A landmark study in lexical ambiguity is Swinney (1979) in which he conducted a series of experiments on the effect of *semantic priming* and lexical access. In his study, Swinney (1979) had subjects listen to audio recordings of two types of utterances: those containing a potentially ambiguous word, such as *bugs* in (3), which could have either the insect sense or the hidden microphone sense, and those in which the ambiguous word had been replaced by an unambiguous one such as *insects* in (4).

- (3) Rumor had it that, for years, the government building had been plagued with problems. The man was not surprised when he found several bugs_◇ in the corner of his room.
- (4) Rumor had it that, for years, the government building had been plagued with problems. The man was not surprised when he found several insects_◇ in the corner of his room.¹

These utterances were then embedded in two types of contexts: those which allowed both interpretations, as in 3, and those which encouraged one interpretation over another (e.g. *The man was not surprised when he found several spiders, roaches and other bugs/insects in his room*). When the recordings reached the potential ambiguity, marked here with (◇), a word would simultaneously appear on a screen placed in front of the subject. The visual word would either be contextually related (e.g. ANT or SPY, depending on the context) or it would be contextually unrelated (e.g. SEW). Subjects were instructed to press a button when they recognised the word on the screen. The time between the word's appearance and the subject's reaction was then recorded.

Swinney (1979) found that, regardless of the context, it would take no less time to recognise a contextually appropriate word, such as ANT, than it would to recognise an inappropriate word SPY. Both of these, however, would be recognised faster than an unrelated

word like SEW. This means that, at least momentarily, all of the possible senses of the ambiguous word were accessed. However, when the visual word was presented three syllables later in the recording, the contextually appropriate words were recognised significantly faster than the inappropriate or unrelated words. By this time, the subjects had already selected their interpretation.

The results from Swinney (1979) have a number of implications not only for how people interpret lexical ambiguity but for language and information processing in general. First, the results show that lexical access is not directly influenced by semantic context in that all of the senses of a word are initially accessed. But soon after they are accessed, context does facilitate the interpretation of the ambiguous word. The speed at which these decisions occur, i.e. within three syllables, lend strong support to the notion that interpretation is incremental. In addition, Swinney proposed that there must be a post-access decision-making process, which is sensitive to context, and selects the contextually-related meanings of the ambiguous word from all of its possible interpretations. These two findings suggest that the process of lexical disambiguation is very rapid and within the span of three syllables, context is said to have an effect on the interpretation of a word. So not only is interpretation incremental, it is context sensitive. This would seem to support the idea that there is at least some level of interaction between the lexical and sentential semantic levels.² One of the lasting importances of Swinney's work has been in controlling textual context, as well as multi-modal information sources during interpretation.

Analysis of Ambiguity in Greenspan Speeches

This section examines how ambiguity is present in Greenspan's speeches. Discourse level ambiguity functions according to the same basic principles as the other types, involving multiple ways of interpreting the same linguistic information, but on a larger scale. We present textual analysis to demonstrate how Greenspan's tendency to equivocate is one form of discourse ambiguity in that conflicting evaluations of the topic at hand open the door to multiple interpretations.

²Another one of the major findings of this study not discussed here is that visual and aural comprehension are linked. What a subject hears can provide a context for what they read. Known as *cross-modal priming*, this suggests that if the mind is indeed modular, there may be an overall interaction between the visual and linguistic processing units.

¹Both of these examples come from Swinney (1979).

Equivocation as Ambiguity While Greenspan's speeches are certainly open to a number of interpretations, it is first necessary to identify the type of ambiguity these speeches exhibit. Most extant work focuses on (sub)sentence level ambiguity or the referential ambiguity that arises through intersentential anaphora. While there may be compulsively ambiguous speakers, just as there exist compulsive liars, there is no evidence of a disproportionate use of those forms of ambiguity in Greenspan's speeches. We claim that the relevant ambiguity in the speeches made by Greenspan must be considered as discourse level ambiguity: equivocation. Equivocation is not self-contradiction: it points to a locus of conflicting information in a form of discourse paraconsistency, leaving it to the interpreter to make decisions based on their own preferences among information sources. For instance, if the subject of the speech were the economy and Greenspan were to say, "The economy has been performing well," followed by "the economy is in trouble," his evaluation of the economy is that it is both good and bad at the same time—a direct contradiction. This does not occur.³ In practice, evaluations such as this are undoubtedly more subtle. For example, in his Humphrey-Hawkins address to the U.S. Senate in February of 1999, Greenspan's evaluation of the current account deficit typifies the type of ambiguity frequently found in his speeches. The first quote warns of the long-term hazards of a growing deficit while the second highlights its short-term benefits.

The rapid widening of the current account deficit has some disquieting aspects, especially when viewed in a longer-term context. Foreigners presumably will not want to raise indefinitely the share of their portfolios in claims on the United States. Should the sustainability of the buildup of our foreign indebtedness come into question, the exchange value of the dollar may well decline, imparting pressures on prices in the United States.

In the recent economic environment, however, the widening of the trade and current account deficits had some beneficial aspects. It provided a safety valve for strong U.S. domestic demand, thereby helping to restrain pressures on U.S. resources. It also cushioned, to some extent, economic weakness in our trading partners.

³Compare this with the opening of *A Tale of Two Cities*, "It was the best of times. It was the worst of times..." Even with a direct contradiction, humans strive for sense and against contradiction.

Thus, the evaluation of one aspect of the economy can be both positive and negative. The basis for the evaluation depends upon the frame of reference against which it is considered, in this case long term versus short term effects of the current account deficit. It is precisely this fragmentation of the corpus of information—or in Lewis's (1982) terms *quarantine*—which allows for the paraconsistency. Incidentally, this is the same "on the one hand, on the other hand" type of equivocation that spurred Harry Truman to wish for a "one-handed economist" (Sicilia & Cruikshank, 2000).

Speeches as Critical Ambiguity Usually the interpretation of ambiguous words or phrases makes little or no difference in the outcome of events. However, sometimes this interpretation can be "critical", having real word consequences. Sicilia and Cruikshank (2000) provide numerous examples of the impact Greenspan's speeches can have on financial markets. In times of trouble, they can provide a much needed boost, as is evident from his September 1998 speech "Question: Is there a new economy?" in which investors detected that he would not raise rates in light of the ongoing Asian financial crisis. Following this speech, the Dow Jones Industrial Average posted a 380 point gain, its largest in a decade. Alternatively, if the market escalates too quickly (in Greenspan's view), they can induce a global sell-off as did his December 1996 speech "The Challenge of Central Banking in a Democratic Society" in which he famously asked: "how do we know when irrational exuberance has unduly escalated asset values?" and markets around the world fell one after the other the next day.⁴

Though not included here for limitations of space, similar analysis these speeches would reveal that Greenspan is capable of indicating a less than equivocal view of the economy. But when he is called upon to make a prediction, one that may have political consequences such as his semi-annual Humphrey-Hawkins address to the U.S. Senate, he is also capable of obscuring any definitive evaluations of the economy, resulting in an ambiguous speech. In the 1999 address quoted above, his evaluation was so mixed that even investors appeared to be unsure how to interpret it. By mid-afternoon, the Dow was down by only 7 points, the S&P 500 index was down less than one point and the dollar was little changed. Such speeches made by

⁴In order, the Japanese Nikkei fell over 3%, Hong Kong dropped 4%, German markets also fell 4%, London's Footsie lost 2% and in the first half hour of trading, the New York Stock Exchange dropped 145 points, triggering the circuit breakers which immediately stop all trading on the market to prevent panic selling.

the chairman of the Federal Reserve and their subsequent impact on financial markets provide a clear case of critical ambiguity.

Having served as chairman of the Federal Reserve and the head of the Federal Open Market Committee (FOMC), the monetary-policy setting branch of the Fed, Greenspan's statements on the economy are naturally of great importance to the financial world.⁵ Traders and investors watch monitor Greenspan's speeches nearly in real time, even making trades during the speech. Those who do not watch the speech, find out what it said soon after. What they monitor for is any hint as to whether the Fed plans to change their rates.⁶ On days following speeches, major newspapers throughout the world analyse the speeches and note any perceived reaction in the markets. It is highly unlikely that in his eighteen years as Fed chief, Greenspan himself has not noticed. Greenspan makes no secret that his speeches are meticulously crafted (Sicilia & Cruikshank, 2000).

While causal connections are frequently drawn in the media between Greenspan's speeches and financial markets, this is often due to the direction and relative size of the reaction. For our purposes, even simply notable reactions will be viewed as interpretations made by a majority of traders in the marketplace.

Summary As mentioned above, it is highly unlikely that Greenspan does not know that investors hang on every word of his speeches looking for clues as to whether he plans to adjust rates or not. Though not included here for limitations of space, similar analysis of his December 1996 speech in which he mentioned "irrational exuberance" and the September 1998 speech in California would reveal that Greenspan is capable of indicating a less than equivocal view of the economy. But when he is called upon to make a prediction, one that may have political consequences, he is also capable of obscuring any definitive evaluations of the economy, resulting in an ambiguous speech. Instances such as the first two examples above do not happen at every Humphrey-Hawkins address by any means. They are simply to illustrate how Greenspan's speeches can occasionally be so ambiguous, not even investors are sure how to interpret them. The third example shows how Greenspan can occa-

sionally make the upside, downside of a trend apply not only to the economy or market in general, but apply it in numerous directions simultaneously.

Again, it is important to mention that this type of ambiguity functions primarily on the discourse level. Accordingly, the experiment reported presently explores the effects of context at that level. If Greenspan presents the upsides and downsides of the economy as an evaluation, knowing that investors are listening for hints as to what he might do in the future, it is possible that the context in which these remarks are interpreted might provide sufficient information to tip the scales in favour of one interpretation or another.

2 *Showing Effects of Non-Local Contexts*

In this section we present an experiment designed to show that manipulating context outside a text can resolve the ambiguity. This idea of this experiment is to create scenarios in which participants act as investor and, faced with contextualizing news of the day (some fabricated) and excerpts from actual Greenspan speeches, they must make financial decisions. The experiment incorporates elements from previous research in ambiguity but rather than focusing on sentence-level disambiguation, implements it at a discourse level. It also involves elements from psychological research such as Kahneman and Tversky's (1979) work in behavioral finance and risk-taking behavior.

Methods and Materials

Assembling the speeches The primary source material for this experiment was a corpus of speeches made by Alan Greenspan.⁷

Greenspan has, of course, made numerous speeches since his appointment in 1987 and not all of them have caused a notable reaction in the stock market. Before incorporating a speech into this study, there had to have been a causal connection proposed between it and the market. While media of various types have made these types of connections, the most readily available sources are newspaper articles. To find the articles, LexisNexis searches were carried out for the keyword "Greenspan" in "Major World Newspapers" between the day of a speech and ending five to seven days after, depending on what day of the week the speech occurred. These searches would result in hundreds of

⁵This paper analyzes his specific impact during his tenure; however, his words since retirement have also been scrutinized in the press.

⁶The FOMC convenes eight times a year to discuss the state of the economy and whether a change in the discount rate, or interest rate that the Fed charges to member banks, is necessary. A change in this rate effects a number of other factors in the economy.

⁷Available here: www.federalreserve.gov — These archives contain speeches dating from 1996 onward; finding earlier speeches proved to be very difficult and was ultimately not pursued.

articles for each speech. Many were reprints from international newswires or articles in smaller regional papers. We concentrated on a select number of reputable international papers: *The New York Times*, *The Washington Post*, *The Guardian*, *The Financial Times*, *The International Herald Tribune*, and occasionally *The Times* (London) and *The Independent* (London). If at least half of these newspapers that printed articles about a speech drew a causal connection, though not necessarily unanimous in their analysis or conclusions, the speech was admitted to the study.

To require each participant to read all twenty-two speeches in their entirety would be too much. Also, it would not accurately reflect how investors obtain their information. During a speech, many will watch the speech live on television and make trades based on what they think the speech indicates from moment to moment. But many more investors will look to financial news services like Bloomberg for a summary of what the speech said, pulling soundbites out of the full text. As Sicilia and Cruikshank (2000) pointed out, the “Irrational Exuberance” speech of 1996 the soundbite factor can also play a role in interpretation. For this reason, the articles that participants read were essentially amalgams of quotes cited in the newspapers as being indicative to investors or accurate summaries of major points of the speech. Minor non-quotative additions had to be made to stitch the quotes together but these were kept as minimally invasive as possible. Usually, the connective remarks were only of the form:⁸

- (5) “It seems persuasive that given the size of the U.S. current account deficit, a diminished appetite for adding to dollar balances must occur at some point,” [the chairman said]. “But when, through what channels, and from what level of the dollar? Regrettably, no answer to those questions is convincing.”

Occasionally, we supplied the topic under discussion:

- (6) [Regarding the recent financial turmoil abroad,] “It is in the interest of the United States and other nations around the world to encourage appropriate policy adjustments and, where required, provide temporary financial assistance,” he said. [But giving the impression that the international

⁸Emphasis is added solely for this description—participants did not have see the inserted text emphasized.

authorities stand ready to guarantee failed domestic businesses,] “could ultimately unbalance the world financial system.”

These were kept as vague as possible while still being true to the content of the speech as to avoid references to specific events which might allow participants to recall the actual context. Finally, since newspapers did not seem to preserve the order of the quotes as they occurred in the speech, only minimal efforts were made to maintain that order here.

Creating the context For the purposes of this experiment, two factors were chosen to represent the context: previous stock market behavior and current events. To show previous market performance, daily stock information was downloaded from the internet⁹ regarding the opening and closing prices of the relevant index, and daily averages were calculated from this information. The graphs used showed the daily averages of the index. For all but one of the scenarios using stock indices the index was the Dow Jones Industrial Average. The one that differed was the NASDAQ Composite. To prevent participants from knowing which index they were seeing, the values from the Y-axis were removed. In addition, the values along the X-axis were modified so that the only reference would be the number of days instead of the actual date.

Headlines placed beneath the graph and before the article were used to give an impression of current events. The majority of these were taken from *The New York Times*’ daily news summaries and business digest pages from the week preceding the speech. Like the analysis and commentary, these were obtained through LexisNexis searches. Headlines related to events that Greenspan explicitly referenced in a speech were given preference. To supplement these, other major economic indicators such as (un)employment figures, inflation data or major commodity price-changes were also included.

If these factors prove to be sufficient context for participants to interpret a speech, their responses should resemble the actual market’s reactions to the speeches. However, merely recreating market behavior is not, by itself, conclusive. In order to test the strength of the contexts, and consequently the degree to which the speeches themselves are ambiguous, nine of the scenarios were given contexts that depicted the opposite from actual events. For the graphs, the X-axis and order of days was reversed so that if the market has been going down for two weeks, it would look as though it had been going up.

⁹The source was finance.yahoo.com (last verified 10 October 2005).

For headlines, any trends in the major economic indicators was likewise reversed. Regarding events explicitly referenced in the speeches, these could not be reversed without directly contradicting the speech but the referent itself could be altered, thus reversing or at least deflecting the perceived impact of the remark. For example, in a speech from September 1998, Greenspan referred to the Asian financial crisis, in which banks and currencies tumbled repeatedly within months of each other, as “vicious cycles”. In the excerpt for the experiment, the explicit mention of Asia was removed so that the text simply read “events abroad”. To complete the fabrication of an alternate context, one of the headlines mentioned a collapse in “Mideast peace talks”, thus creating a situation in which the suggested reference could be interpreted as a situation of political conflict instead of economic turmoil.

Presentation of the materials The materials consisted of twenty-two single-page scenarios containing a graph,¹⁰ several headlines of the day, and a short article gisting one of the speeches collected. Beneath each article, participants were instructed:

- (7) Your current holdings consist primarily of shares in well-established industrial producers but also some newer technology firms. The graph at the top of this page indicates the recent overall performance of these investments. Based on the article you have just read, would you...

Participants were then presented with three options: “Buy”, “Sell” or “Maintain your current holdings.” The description of the portfolio was modified in each case to approximate the sector of the market in which the newspapers reported a reaction. Thus, if it was noted that technology stocks fell, the participant was told that the majority of their investments were in technology firms or high-tech producers. Since the speeches date from the mid-1990’s onward, which saw the boom and bust cycle of technology firms in the United States, many of the holdings involved technology-oriented shares. Instructions were phrased in as neutral a manner as possible to avoid encouraging one choice over another and differed only in the description of the portfolio. They were presented in random order to reduce any effect order might have on participants’ decisions.

¹⁰Twenty-one involved stock indices (twenty showing the Dow Jones and one showing the NASDAQ), and one pertained to currency markets (the U.S. Dollar against the Japanese Yen). Since currency graphs depict the price similar to a stock graph, these differences in the indices did not pose a problem.

Collecting the data In the interest of extending the reach of the participant pool as wide as possible, the experiment was also to be made available online. This choice was made with awareness of the advantages and risks associated with running experiments over the internet (Hewson, Yule, Laurent, & Vogel, 2003), using a system developed by the Computational Linguistics Group at Trinity College Dublin (Buckley, 2004; Graham, 2005). We adopted an arbitrary (as opposed to random or fully targeted) sampling strategy. We began posting notices first on the TCD electronic notice-board and then on various financially-oriented newsgroups through Google. Several factors taken into consideration in selecting groups. Since the majority of participants already obtained were not professional investors, we wanted to attract as many people who had a familiarity with investing as possible. The first consideration was the focus of the newsgroup. Only groups with a specified focus on investing, economics or finance were considered. The second consideration was the amount of traffic on the newsgroup. Obviously, the higher the traffic, the more people would receive information about the experiment, but upon examination the newsgroups with the highest traffic were usually dominated by advertisements and solicitations and showed no evidence of interaction between members. For this reason, we considered only those which appeared to have ongoing conversations updated on a regular basis. These groups turned out to be harder to find and lower in active members than expected. Most groups were ranked as “low” traffic and had between ten and 100 members. In all, we posted to eighteen groups over a period of ten days and would periodically check to see if people had responded within the newsgroup.

Between the posting on the TCD notice-board and the Google groups, we obtained a total of five more participants, three of them from Trinity. We also solicited participation from students and staff of the Business, Economics and Social Studies (BESS) programme at Trinity and the Stern Business School in New York University. Because of the nature of the sampling method it isn’t possible to calculate a response rate. The actual participation rate was quite low.

The results of this experiment were taken from the eighteen actual participants. There were ten male and eight female. Nine participants were between 21 and 30, two were between 31 and 40, three were between 41 and 50 and four were between 51 and 60. Six people considered themselves to be “Not at all familiar” with investing in financial markets, five consid-

ered themselves “Somewhat familiar”, six people considered themselves to be “Fairly familiar”, and one person considered himself to be “Very familiar”. Of the eighteen participants, two had professions that involved investing in, analysing or otherwise dealing with financial markets.

Results

In discussing these results, it is first necessary to define some basic terminology used to describe the aspects of a scenario in comparison. For the purposes of analysis, the reactions to Greenspan’s speeches were categorised as either “buying” or “selling”, even if every buying event is also a selling event. If the market went up following a speech, this was considered a “buying” reaction and if it went down, this was considered a “selling” reaction. The scenarios in the experiment were likewise categorised in a similar manner. This terminology naturally presents a naïve perspective of the market. Nevertheless, it will suffice as a method of describing the reactions that occurred and comparing them with the results obtained.

It was also necessary to distinguish between those presented with “real” contexts and those in which the context was “fabricated”. It is important to note that the buy/sell nature of a given scenario is subject to the context. Thus, if the market “bought” after a particular speech and is here presented with a fabricated context, it would then be categorised as a “selling” scenario. The reverse is also true that a “selling” scenario in a real context would be considered as a “buying” scenario in a fabricated one.

For the purposes of discussion, decisions will be referred to here as either “correct” or “incorrect” with respect to their scenario. Decisions will be termed “correct” if they correspond to the buy/sell category of the scenario and “incorrect” if they do not. For example, in a buying scenario presented with a real context, a decision to “buy” would be termed correct and “sell” incorrect. If the context of this scenario had been fabricated, these values would reverse so that “sell” would be correct and “buy” incorrect. Decisions to “maintain” in any scenario were not considered as either correct or incorrect but rather as refraining from making a decision. The frequency of such decisions touches on a broader discussion of the psychology of risk-aversion behaviour, particularly with regard to financial decision-making and is discussed below with reference to work by Kahneman and Tversky (1979).

A comparison of decisions to buy, sell, or maintain in buying scenarios and selling scenarios did reveal a significant difference ($\chi^2 = 3D 11.052$, $df = 3D 2$,

$p \leq 0.01$). This indicates that participants had clear preferences to choose one option over another depending on the scenario. Essentially, in buying scenarios people generally bought and in selling scenarios they generally sold (see Table 1). This means that participants were guided in their interpretation by context: where definite decisions to buy or sell were made, they tended in the direction suggested by context even though the textual material of the speech remained unchanged ($\chi^2 = 3D 10.661$, $df = 3D 1$, $p \leq 0.01$).

Table 1: Decisions in Buying v. Selling Situations

	Buy	Sell	Maintain	Total
Buying Situation	60	44	76	180
Selling Situation	48	84	84	216
Total	108	128	160	396

A comparison of decisions to buy, sell, or maintain between real and fabricated contexts did not reveal a significant difference ($\chi^2 = 3D 2.551$, $df = 3D 2$, $p \leq 1$). This shows that participants were not able to distinguish the real contexts from the fabricated ones indicating that, regardless of whether the context was real or not, their decisions were unaffected (see Table 2).

Table 2: Decisions in Real v. Fabricated Contexts

	Buy	Sell	Maintain	Total
Real Context	57	80	97	234
Fabricated Context	51	48	63	162
Total	108	128	160	396

Likewise, a comparison of risk-aversion (i.e. people who bought or sold versus those who maintained) in real and fabricated contexts did not reveal a significant difference ($\chi^2 = 3D 0.576$, $df = 3D 1$, $p \leq 1$). Thus, people were no more or less likely to make a decision in a real context than a fabricated one (see Table 3).

Table 3: Decisions to Maintain in Real v. Fabricated Contexts

	Buy or Sell	Maintain	Total
Real Context	137	97	234
Fabricated Context	99	63	162
Total	236	160	396

Similarly, the accuracy of the decision (i.e. correct versus incorrect answers) in real and fabricated contexts was not significantly different ($\chi^2 = 3D 0.2773$, $df = 3D 2$, $p \leq 1$), suggesting that participants were no more or less likely to make the correct decision in real contexts than in fabricated ones (see Table 4).

Table 4: Accuracy in Real v. Fabricated Contexts

	Correct	Incorrect	Maintain	Total
Real Context	80	57	97	234
Fabricated Context	57	42	63	162
Total	137	99	160	396

These comparisons between real and fabricated contexts show that the “reality” of the context presented did not affect participants’ decisions to buy or sell, their willingness to make such a decision, or the accuracy with which they made it. In light of these encouraging results, their decisions should be equally reliable in both buying and selling scenarios.

The risk-aversion comparison (i.e. the willingness to make a decision versus maintain) was not significant ($\chi^2 = 3D 0.453$, $df = 3D 1$, $p \leq 1$), suggesting that participants were again no more or less likely to refrain from making a decision whether they were in a buying scenario or a selling scenario (see Table 5). The

Table 5: Decisions to Maintain in Buying v. Selling Situations

	Maintain	Buy or Sell	Total
Buying Situations	76	104	180
Selling Situations	84	132	216
Total	236	160	396

difference between the number of correct responses in buying scenarios versus and selling scenarios was also not also significant ($\chi^2 = 3D 1.312$, $df = 3D 2$, $p \leq 1$), suggesting that participants were no more or less likely to make the correct decision in a buying situation than a selling situation (see Table 6). Taken together, these results indicate that in buying and selling scenarios, participants were inclined to either buy or sell and this inclination was not significantly affected by the reality of the context presented. Risk-aversion behaviour was not affected by the buy/sell category of the scenario or by the reality of the context. Finally, the accuracy with which these decisions were made

Table 6: Accuracy in Buying v. Selling Situations

	Correct	Incorrect	Maintain	Total
Buying Situations	60	44	76	180
Selling Situations	84	48	84	216
Total	144	92	160	396

was also not affected by the buy/sell category or the reality of the context. In sum, these results suggest that in any given scenario, the decisions made by participants were made with equal willingness to select one option or the other, to act decisively or refrain, and with the same level of accuracy.

Discussion

It is obvious that other experimental manipulations would be illuminating. For example, we have an effect of nonlocal context type on resolution of equivocation, but we have not constructed a context vs. no-context condition. More procedurally, revised experiments must improve on response rates, and one potential factor in that is overall complexity of the materials. Another potential issue is the overall familiarity of the participants with investing strategies. As mentioned above, only two out of eighteen participants had professions that involved regular contact with financial markets. Perhaps if the participant pool had incorporated more financial professionals, who would be accustomed to interpreting the type of language and data presented here, the results would have been crisper. Revising the sampling strategy along those lines might have an impact on the perception of complexity within the materials.

Nonetheless, the results from this experiment are interesting on a number of levels. First, the evidence that participants had clear preferences toward either buying or selling in the appropriate scenario reinforces the nature of the categories themselves. This coupled with the lack of any significant difference between these preferences with respect to the context suggests that participants were unable to distinguish real scenarios from fabricated ones. These findings are also bolstered by the lack of any significant difference in the willingness to make a decision in any scenario. If neither the category of the scenario nor the “reality” of the contexts affected the participants’ decisions, it can be said with reasonable certainty that the decisions made in this experiment were made with equal

reliability.

These findings also support several theoretical approaches to the interaction between context and interpretation. Most directly, they speak to Hirst's (1987) claim that interpretation is to a large extent determined by context. While his comments directly addressed issues of lexical and syntactic ambiguity, there does seem to be an indication here that this may be the case. Moreover, we have provided concrete empirical support to back up the commonly expressed intuition. Context was able to influence the interpretation significantly in this experiment, in some cases actually reversing the interpretation of the speech. Also relevant is Poesio's (1996) distinction between semantic ambiguity and perceived ambiguity in that the interpretations here seem to have been made with regard to only the contextually available information instead of the full range of semantic interpretation. This, of course, assumes that the ambiguity was to some degree perceived in the experimental materials.

Also relevant to this study is the risk-aversion factor which has been studied in great detail following the work of Prospect Theory and its application to behavioural finance (Kahneman & Tversky, 1979). Briefly stated, Kahneman and Tversky (1979) found that when subjects were presented with a sum of money and then given the opportunity to increase a sum of money by half with 100% certainty or double their sum with only a 50% chance and risk losing everything, the majority of people will take the more certain option. Thus, the prospect of loss is perceived as greater than the prospect of gain. In keeping with the findings of Prospect Theory, all of the scenarios in this study offered the option to "Maintain your current holdings" and refrain from making a decision. In every scenario at least four people and, in a few scenarios, the overwhelming majority selected this option. This might be indicative of similar risk-aversion behaviour but it could also indicate an inability to make a satisfactory interpretation of the speech.

Two other issues highlighted by Kahneman and Tversky (1979) are the problematic nature of hypothetical decision-making and the cumulative effect of gain or loss. The caveat regarding the former is that people might be more (or perhaps less) risk-seeking in hypothetical situations than they would in real life. The effect of the hypothetical nature of this experiment on the results obtained cannot be quantified. Regarding the latter, Kahneman and Tversky (1979) suggests that participants who "have not made peace with their losses"¹¹ become less risk averse with each

decision, taking risks that they otherwise would not take. Earlier formats of this experiment tried to include elements such as a budget or a reveal panel, as mentioned in Section 6, which might have been able to monitor this effect but as participants were only asked to make decisions and they were shown no result from any decision they made, there can be no actual cumulative perception of gain or loss in this study.

3 Conclusion

We have described an experiment whose results demonstrate the effect of non-local context on resolving discourse ambiguity. Although the participant numbers are small, the results appear robust: they are consistent with other empirical study of risk aversion; and they provide empirical support for the intuition that global contexts can resolve ambiguity. This claim is typically assumed without proof. We have not experimented with context vs. no-context, but have shown a clear pattern in the results by manipulating context in favor of one interpretation or another. We conjecture that risk aversion strategies would be more prevalent in the absence of context altogether. The context of the scenarios presented with the speeches had a strong influence on the interpretation, in some cases reversing the decisions made by the actual stock market. These results provide support for some previous theories on the effects of context on interpretation (Hirst, 1987; Poesio, 1996) and demonstrates that these issues do operate on a discourse level. There also may be a connection in this experiment to previous research in risk aversion and behavioural finance (Kahneman & Tversky, 1979). Although this particular version did not attempt to quantify this element, the noticeable number of participants that refrained from making decisions on a given question suggest that this may be worth pursuing.

References

- Altmann, G. T., Garnham, A., & Henstra, J.-A. (1994). Effects of Syntax in Human Sentence Parsing: Evidence against a structure-based proposal mechanism. *Journal of Experimental Psychology*, 20(1), 209–216.
- Buckley, M. (2004). Cross-disciplinary Research into improving Internet-based Research Methodology Applied to the Interpretation of Natural Language Quantifiers. Master's thesis, Compu-

¹¹(Kahneman & Tversky, 1979, p. 288).

tational Linguistics Group, Trinity College, University of Dublin.

- Graham, Y. (2005). Services for Experimentation in the Human Sciences: An Online Experimentation Tool. Master's thesis, Computational Linguistics Group, Trinity College, University of Dublin. Under Examination.
- Hewson, C., Yule, P., Laurent, D., & Vogel, C. (2003). *Internet Research Methods: A practical guide for the social and behavioural sciences*. Sage Publications: London.
- Hirst, G. (1987). *Semantic Interpretation and the Resolution of Ambiguity*. Cambridge University Press: London.
- Kahneman, D. & Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292.
- Lewis, D. (1982). Logic for Equivocators. *Noûs*, 16, 431–441.
- Poesio, M. (1996). Semantic ambiguity and perceived ambiguity. In van Deemter, K. & Peters, S. (Eds.), *Semantic Ambiguity and Underspecification*, pp. 159–196. CSLI: Stanford, CA.
- Sicilia, D. B. & Cruikshank, J. L. (2000). *The Greenspan Effect: Words that move world markets*. McGraw-Hill: London.
- Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)Consideration of contextual effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 645–659.