Published in final edited form as:

Neuropsychologia. 2012 April ; 50(5): 674–687. doi:10.1016/j.neuropsychologia.2012.01.004.

fMRI evidence for strategic decision-making during resolution of pronoun reference

Corey T. McMillan^{1,*}, Robin Clark², Delani Gunawardena¹, Neville Ryant², and Murray Grossman^{1,*}

¹University of Pennsylvania School of Medicine, Department of Neurology

²University of Pennsylvania, Department of Linguistics

Abstract

Pronouns are extraordinarily common in daily language yet little is known about the neural mechanisms that support decisions about pronoun reference. We propose a large-scale neural network for resolving pronoun reference that consists of two components. First, a core language network in peri-Sylvian cortex supports syntactic and semantic resources for interpreting pronoun meaning in sentences. Second, a frontal-parietal network that supports strategic decision-making is recruited to support probabilistic and risk-related components of resolving a pronoun's referent. In an fMRI study of healthy young adults, we observed activation of left inferior frontal and superior temporal cortex, consistent with a language network. We also observed activation of brain regions not associated with traditional language areas. By manipulating the context of the pronoun, we were able to demonstrate recruitment of dorsolateral prefrontal cortex during probabilistic evaluation of a pronoun's reference, and orbital frontal activation when a pronoun must adopt a risky referent. Together, these findings are consistent with a two-component model for resolving a pronoun's reference that includes neuroanatomic regions supporting core linguistic and decision-making mechanisms.

Keywords

decision-making; language processing; pronouns; fMRI; lexical semantic

Personal pronouns like “he” and “she” are extraordinarily frequent in daily language. They refer to a previously mentioned or acknowledged person during discourse, but pronouns themselves carry little meaning. For example, in the sentence “He cried,” we know that a male cried but we do not know exactly which male cried. Given the frequency of pronouns in everyday language, it is important for models of language processing to be able to account for how individuals resolve a pronoun's reference. Moreover, a long tradition of psycholinguistic research has investigated how individuals determine a pronoun's meaning, but relatively little is known about the neural mechanisms that support this process. In this study, we examine the neural basis for the linguistic and decision-making processes that contribute to establishing pronoun reference.

© 2012 Elsevier Ltd. All rights reserved.

*Corresponding Authors: University of Pennsylvania Medical Center, Department of Neurology, 3400 Spruce Street, 3 West Gates, Philadelphia, PA 19104, mcmillac@mail.med.upenn.edu, mgrossma@mail.med.upenn.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A major challenge for models of pronoun resolution is that a pronoun is by nature somewhat ambiguous since the pronoun itself has minimal meaning, and instead its meaning must be derived from a discourse referent. A wealth of psycholinguistic studies has investigated how readers determine a pronoun's referent (Almor, 1999; Caramazza, Grober, Garvey, & Yates, 1977; Crawley, Stevenson, & Kleinman, 1990; Garnham, 2010; Garnham, Oakhill, & Cruttenden, 1992a; Garvey & Caramazza, 1975; Kennison & Trofe, 2003; Sanford, 1982). From these studies it has become clear that several types of linguistic cues may be used to determine a pronoun's referent including, but not limited to: verb-bias, focus, gender, implicit causality, grammatical roles, and subjecthood (e.g., general preference for a sentence's subject; see Kehler, Kertz, Rohde, & Elman (2008) for a review of different linguistic cues). Consider, for example, the following sentences:

1. The grandmother hugged the uncle. He smiled.
2. The grandfather hugged the uncle. He smiled.

In sentences like (1) noun gender information can be used to help identify the pronoun's referent since there is only one male and therefore "he" must refer to the "uncle". However, in sentences like (2) gender is not informative because both potential referents are the same gender as the pronoun. Another cue that can be used to determine a pronoun's referent is the verb's semantic meaning, which suggests that a given process or state (e.g., being hugged) is the cause for another (e.g., smiling). For example, while gender information is not informative in (2), a reader can use the semantic information of the verb to generate an inference - since the uncle received a hug, he is more likely to have smiled. In this paper we focus on gender information and verb-bias information, two types of cues that have been well-documented in the psycholinguistic literature.

Gender information is well known to contribute to the interpretation of a pronoun's referent (Kennison & Trofe, 2003). For example, Kennison & Trofe (2003) presented readers pairs of sentences containing a gender-stereotyped noun and observed significantly longer reading times when the gender-stereotype of the noun mismatched the gender of the pronoun (e.g., "The executive distributed a memo. She made it clear that work would continue as normal."). This finding demonstrates that when readers encounter a pronoun like "she" they attempt to resolve its meaning by assigning it to a female antecedent and encounter difficulty when the antecedent is less consistent with the expected gender.

It has also been demonstrated that verb-bias information contributes to sentence comprehension in general (Pickering & Majid, 2007; Rudolph & Försterling, 1997) and more specifically to the interpretation of a pronoun's referent (Garnham et al., 1992a; Garnham, Traxler, Oakhill, & Gernsbacher, 1996b; Garvey & Caramazza, 1975). For example, in a seminal study Garvey & Caramazza (1975) presented participants with pairs of sentence fragments containing an ambiguous pronoun that cannot be assigned to an antecedent noun based solely on gender (e.g., "The father confessed to his son because he..." and "The father scolded his son because he..."), and asked participants to complete the sentences. They observed a strong bias in the interpretation of "he" referring to either the subject noun ("father") for the verb "confessed" or the object noun ("son") for the verb "scolded". This finding established empirical evidence that information from the verb alone biased the reader's interpretation of a pronoun.

Little is known about the neuroanatomic mechanisms that support the relative uses of gender and verb-bias information in the resolution of a pronoun's reference. On the one hand, a small number of fMRI studies have emphasized the contribution of left peri-Sylvian regions in the posterior-inferior frontal lobe and the posterior-superior temporal lobe to core linguistic processes involved in pronoun resolution such as the gender associated with a noun or the grammatical roles associated with a verb in a sentence (Hammer, Goebel,

Schwarzbach, Münte, & Jansma, 2007). On the other hand, these studies appear to implicate additional neuroanatomic regions beyond those immediately entailed by linguistic properties like the semantic information such as gender and grammatical information such as a verb's bias. In one fMRI study investigating the role of gender cues during German pronoun resolution, Hammer et al. (2007) compared sentences containing pronouns that have a congruent structure with a matching noun and pronoun gender (e.g., "The woman is popular because she is beautiful.") and an incongruent structure due to a mismatch between noun and pronoun gender (e.g., "The woman is popular because he is beautiful."). Relative to a fixation baseline, they observed frontal cortex activation, including bilateral inferior (Brodmann Areas (BA) 45, 44), left dorsolateral prefrontal cortex (dlPFC; BA 9) and bilateral dorsal inferior (BA 6) prefrontal regions, as well as left temporal cortex, including BA 38, 22, and 21. However, because this analysis included incongruent items with semantic violations and had a weak baseline, it is not clear which of these regions is contributing to the language-specific mechanisms that support successful resolution of a pronoun's referent and which regions are contributing to attentional mechanisms that support the detection of a semantic violation.

Nieuwland et al. (2007) directly compared different kinds of coherent Dutch sentences that contain a pronoun. They observed that sentences with referential ambiguity due to the absence of gender information cues (e.g., "Ronald told Frank that *he* had a positive attitude towards life.") recruit right superior (BA 8/9) and medial frontal (BA 10) cortex along with bilateral inferior parietal (BA 39) and medial parietal (BA 7/31) cortex relative to unambiguous sentences in which gender cues are informative (e.g. "Ronald told Emily that *he* had a positive attitude towards life."). These regions are not typically associated with language-specific processing but are reported in the decision-making literature. This prompted Nieuwland et al. (2007) to conclude that, beyond strictly linguistic processes like gender and verb-bias, decision-making resources also may contribute to resolving a pronoun's referent. The investigations by Hammer et al. (2007) and Nieuwland et al. (2007) thus both provide evidence that anterior prefrontal and parietal cortex are recruited to support pronoun resolution, but the studies were not designed to specifically investigate the role of non-language regions commonly associated with decision-making.

The present study evaluated the role of decision-making in resolving a pronoun's referent. In our approach to decision-making, we adopted a game-theoretic model of reference resolution (Clark, 2011; Clark & Parikh, 2007). According to Clark's (2011) model, humans behave as rational agents when using language. This does not imply that humans are infallible in their language use, but merely that humans use language to accomplish the common goal of expressing meaning to one another as clearly as possible while minimizing computational processing demands. Specifically, Clark & Parikh (2007) proposed that readers use two mechanisms in order to correctly interpret a sentence. First, readers strategically choose a pronoun's referent in a probabilistic manner that maximizes the likelihood of correctly interpreting a sentence. Second, when readers interpret a sentence, they compute the relative value associated with the "risk" of choosing the incorrect referent, and the computational resources required to minimize the risk of misinterpreting a pronoun's meaning. We provide a detailed discussion and examples of our proposed game-theoretic model in the Appendix.

A game-theoretic framework has three advantages over extant psycholinguistic models of pronoun resolution. First, while psycholinguistic investigations have determined which linguistic cues are used to determine a pronoun's referent, they do not currently account for the process of how readers strategically choose one cue over another cue. A game-theoretic model allows the investigator to quantify (and predict) the relative role of linguistic cues depending on the probability that a cue is informative (e.g., *Is gender information useful in*

this context?), the “riskiness” of using a particular cue (e.g., *Is assigning “he” to a gender-neutral referent like “child” likely to result in a misinterpretation?*) and the computational resources required to evaluate each cue (e.g., *Is it worth the time and resources required to consider all of the potential cues if gender information alone can yield a correct interpretation?*). A second, resource-related advantage of a game-theoretic approach may be related to its ability to inform evolution of language, such as the origins of pronoun use. Pronouns, for example, may have been adopted over time because they enable speakers to express meaning to an interlocutor with a very small number of words that are easy to retrieve because they are highly frequent and can be used in most circumstances while minimizing the resources required to repeat a much less frequent noun phrase that is useful in a very limited number of circumstances. Finally, game theory provides a universal framework that quantifies several domains of human behavior, including but not limited to financial decision-making (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003), learning (Camerer, 2003), social preferences (Camerer, 2003), mate selection (Miller & Todd, 1998) and psychiatric disorders (Patokos, in press). From this perspective, there is no *a priori* motivation for hypothesizing that decision-making during language processing is in some way distinct from other domains of human behavior, and we therefore hypothesize that similar decision-making mechanisms support language processing. Likewise, we emphasize that our hypothesized decision-making mechanism is not necessarily explicit, resembling decision-making in other cognitive domains.

From this perspective, we hypothesize that the probabilistic and risk-related decision-making mechanisms proposed by Clark & Parikh (2007) have distinct neural substrates situated in a frontal-parietal decision-making network (see below), and that these mechanisms interact with core-language processing mechanisms in peri-Sylvian cortex that support the semantic and syntactic resources that help compute linguistic aspects of pronoun meaning. In sum, we hypothesize that both linguistic processes and decision-making processes contribute to resolving a pronoun’s referent, and that a large-scale neural network supports this by recruiting both peri-Sylvian regions related to linguistic processes and frontal-parietal regions related to decision-making. This is consistent with a model of the neural basis for language processing that involves a core language processing mechanism in left peri-Sylvian cortices that recruits additional executive resources, as needed, to support the interpretation of sentence-level material (Cooke et al., 2006; Novais-Santos et al., 2007). We describe each of these hypothesized components in more detail in the following sections.

Core language resources for resolving a pronoun’s referent

One component of our hypothesized model includes a core language processing network that supports processing associated with the linguistic attributes of a pronoun. On the one hand, lexical semantic resources are required in order to interpret the gender attributes of a pronoun and a potential referent. For example, “she” can refer to “nun” but does not ordinarily refer to “king” (here and throughout the paper, we do not consider non-standard uses of pronouns for metaphor and other pragmatic considerations). fMRI studies investigating grammatical gender have revealed left inferior frontal (BA 44/45) as well as middle and superior temporal cortex (BA 21/22) recruitment when individuals must make explicit gender judgments during language comprehension (Miceli et al., 2002) and language production (Heim, Opitz, & Friederici, 2002; see Heim, 2008 for a review). Indefrey & Levelt (2004) have argued that middle temporal cortex is involved in the selection of a word form and this selection process may include accessing gender information. When individuals must access gender information of a noun in order to resolve a referent, we predict that we will observe middle and superior temporal cortex activation to support lexical semantic resources required for the retrieval of gender information.

Beyond lexical semantic resources required to determine gender, we argue that additional neural resources are required to determine the subtle bias of a verb toward emphasizing its subject or object, that is, the verb-bias. Since verb-bias information may involve selecting between two potential referents of a pronoun that are serving different grammatical roles in a sentence (e.g. subject or object), we may also observe activation of inferior frontal cortex (BA 44/45), a region that has commonly been implicated as supporting syntactic resources during sentence comprehension (see Grodzinsky & Santi, 2008 and Kaan & Swaab, 2002 for recent reviews). For example, several fMRI studies have reported increased inferior frontal activation for object-initial sentences (Caplan, Alpert, Waters, & Olivieri, 2000; Constable et al., 2004; Cooke et al., 2002; Just, Carpenter, Keller, Eddy, & Thulborn, 1996), and this may be related to increased processing demands associated with computing a syntactic structure when a non-canonical (object-subject) word order is encountered. In a combined fMRI and lesion-patient study, Tyler et al. (2011) observed recruitment of inferior frontal cortex and insula associated with increased syntactic demands due to a verb's bias toward a linguistic structure. Moreover, they observed a positive correlation between inferior frontal and insula activation and patients' acceptability judgments of sentences in which a higher acceptability judgment indicated an increased sensitivity to syntactic information. Following from these findings, we predict that when lexical semantic gender information can not be used to choose between one referent or another in sentences such as (2), there will be increased demands on verb-bias information and thus increased inferior frontal and insula activation.

Decision-making resources for resolving a pronoun's referent

Another component of our model proposes that strategic decision-making resources contribute to the resolution of pronoun reference. Specifically, we propose that a probabilistic mechanism plays a role in determining the likelihood of a pronoun referring to a particular referent. For example, the gender-neutral noun "child" could be either a female or male, and therefore it is necessary to evaluate the probabilistic likelihood of the association of the noun with each gender when we encounter nouns with ambiguous gender associations. We predict the highest probabilistic demands during pronoun reference resolution when determining the referent from a sentence containing two gender-neutral nouns such as "The teacher billed the patient. He groaned." This is because this discourse requires individuals to evaluate the probability of the gender of both the subject (e.g., "teacher") and object (e.g., "patient") of the sentence.

We hypothesize that dlPFC will be recruited to support a probabilistic mechanism. Neuroimaging investigations have directly associated dlPFC with calculating target probabilities during decision-making tasks (Casey et al., 2001; Scheibe, Ullsperger, Sommer, & Heekeren, 2010) and contributing to probability calculations of fairness in the Ultimatum Game, an economic game that requires individuals to decide whether to accept or reject a financial offer (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003; see also Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006 for convergent repetitive transcranial magnetic stimulation evidence). Nonhuman primate investigations provide additional support for the role of dlPFC as a probabilistic mechanism (Barraclough, Conroy, & Lee, 2004; Kim & Shadlen, 1999; Leon & Shadlen, 1999). For example, Leon & Shadlen (1999) demonstrated that neurons in BA 46 fire at a magnitude proportional to an expected reward in a motion discrimination task.

A second decision-making component may function as a value mechanism to support strategic decision-making during pronoun reference resolution. Value in this context refers to the "risk" associated with choosing an incorrect referent and therefore misinterpreting a sentence. In daily language, individuals must often choose a referent with some uncertainty

because the context does not allow the reader to directly determine a mapping between pronoun and referent. For example, in the discourse “The grandfather hugged the child. She smiled.” it can be determined that “she” does not refer to “grandfather”. Thus the reader can indirectly determine that “she” must refer to “child.” However, because “child” could be either a female or a male, assigning “she” to “child” has some risk associated with it.

We are not aware of any empirical investigations directly considering the role of risk in language interpretation (cf. (Clark, 2011; Pinker, Nowak, & Lee, 2008; Sally, 2003). However, as stated previously, we hypothesize that the same mechanisms that have been proposed to support risk in a game-theoretic framework in other domains of human behavior (e.g., financial decision-making) will be recruited to support risk in the context of language processing. Neuroimaging investigations of explicit decision-making in non-linguistic tasks such as gambling tasks have frequently implicated orbital frontal cortex (OFC; BA 11) for value assessment (Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Fiddick, Spampinato, & Grafman, 2005; Hsu, Bhatt, Adolphs, Tranel, & Camerer, 2005; Sanfey, Loewenstein, McClure, & Cohen, 2006). Converging evidence comes from patient investigations demonstrating that OFC disease significantly impairs value-based decisions (Adolphs, 1999; Rogers et al., 1999; Rolls, 2000; Rosen et al., 2005; Salmon et al., 2003). Based on these considerations, we predict OFC activation in risky pronoun referent assignments, such as that associated with assigning a gender-specific pronoun to a gender-neutral noun.

A final component of this decision-making model is hypothesized to integrate probabilistic and value information into a single currency, “expected utility” (EU). Formally, EU is equal to the probability of an outcome times the value or risk associated with choosing an outcome. According to Clark and Parikh’s (2007) game theoretic account for pronoun reference resolution, individuals use EU to strategically make a decision about a pronoun’s referent in a manner that maximizes the likelihood of choosing the correct referent while minimizing the resources required.

Neuroimaging studies implicate inferior parietal cortex (IPC) in the integration of the components contributing to EU (Huettel, Song, & McCarthy, 2005; McClure, Laibson, Loewenstein, & Cohen, 2004; Paulus et al., 2001; Platt & Glimcher, 1999; Schall, 2001; Vickery & Jiang, 2009). For example, Vickery & Jiang (2009) reported increased IPC activation in a “matching-pennies” game when individuals made a forced-choice decision (choosing to move a soccer goalkeeper left or right) to predict a computer’s choice (kicking ball left or right). Activation in IPC was greater when participants were given feedback than when not, and was greater in uncertain conditions compared to certain conditions. This suggests that IPC activation is sensitive to both reward information and probabilistic information, respectively. Converging evidence for the role of IPC in the integration of information to form a decision comes from several primate studies. These studies consistently demonstrate that neural activation in parietal cortex scales proportionally with EU (Glimcher, Dorris, & Bayer, 2005; Gold & Shadlen, 2001; Kiani & Shadlen, 2009; Shadlen & Newsome, 2001), a phenomenon labeled “physiological expected utility” (Glimcher et al., 2005). Thus, we also predict IPC activation as part of the decision-making component during resolution of a pronoun’s referent.

In the current study we investigated the proposed two-component model using fMRI. We presented readers with pairs of grammatically simple sentences: one sentence containing two nouns (e.g., “The boy fed the grandfather.”) and a simple sentence containing a pronoun (e.g., “He grinned”). We monitored brain activation while participants chose the pronoun’s referent. We manipulated the lexical semantic bias of each noun by using combinations of nouns that are definitionally male or female (e.g., grandfather or woman) or are gender-neutral (e.g., child), according to pretest norms. We predicted that individuals would recruit

cortical regions supporting core linguistic processes in inferior frontal and middle temporal cortex, and regions supporting decision-making in dlPFC, OFC, and IPC. Through manipulations of the nouns and pronouns in the stimulus materials detailed in the Methods section, we sought to demonstrate the mechanistic contribution of these brain areas to resolving a pronoun's reference in more detail.

Methods

Participants

18 healthy young adults from the University of Pennsylvania community participated in the study for monetary payment. All participants were native speakers of English, right-handed, and in good health with no history of neurological difficulty. Informed consent was obtained from all participants according to a protocol approved by the University of Pennsylvania Institutional Review Board. We excluded two participants due to excessive motion artefact in the MRI scanner (>3mm movement in any direction) and therefore all analyses reported are for 16 healthy adults.

Behavioral Stimuli

We identified 80 nouns, half of which were gender-neutral and half gender-biased. To assess the gender of each noun, we asked 20 college-age volunteers to rate each noun (e.g., “woman”, “grandfather”, “child”) on a 5-point scale, where a 1 represented definitely-female, a 5 represented definitely-male, and the points in between represented weaker degrees of a gender-bias with 3 equally likely to be a male or female. We then converted these scores from a 5-point scale ranging from female-bias to male-bias to a 3-point scale which represents the degree of gender bias for each noun independent of a particular gender: a “1” represented no gender-bias (equal to a value of “3” from the 5-point scale representing no gender-bias), a “2” represented a weak gender-bias (equal to a value of “2” from the 5-point scale indicating a weak female-bias or a value of “4” from the 5-point scale indicating a weak male-bias), and a “3” represented a strong gender-bias (equal to a “1” out of 5 points indicating a strong female-bias or a “5” out of 5 points indicating a strong male-bias). A statistical analysis confirmed that gender-biased nouns ($M = 2.91$, $SD = 0.15$) are significantly more biased than gender-neutral nouns [$M = 1.12$, $SD = 0.12$; $t(78) = 59.01$, $p = 0.000$]. Additionally the gender-biased and gender-neutral nouns were matched for written lexical frequency [$t(78) = 0.79$], familiarity [$t(78) = 1.46$], and imageability [$t(78) = 1.21$].

Using these nouns, a total of 200 sentences were generated by pairing male-biased nouns, female-biased nouns, and/or gender-neutral nouns with a verb to create coherent and meaningful experimental stimuli (e.g., “The grandmother hugged the groom”). Each noun was repeated a maximum of six times and each noun-noun pair was never repeated. Half of the verbs were alternating-dative (e.g., called) and half reciprocal (e.g., kissed) verbs. The sentences generated from these nouns comprised the following conditions:

- *Directly-Determined* (e.g., “The woman paid the boy. He pouted”): contained 1 female-biased noun and 1 male-biased noun, and the pronoun can refer directly only to one of the preceding referents.
- *Indirectly Determined* (e.g., “The mom served the child. He pouted”): contained 1 gender-neutral noun and 1 gender-biased noun. The gender of the pronoun does not match the gender of the gender-biased noun, thus the pronoun cannot refer to the gender-biased noun and therefore indirectly refers to the gender-neutral noun.

- *Undetermined*: the pronoun can be assigned to the subject or the object of the preceding sentence. We further manipulate these noun pairs into two different types of combinations:
 - Gender-Biased (e.g., “The boy fed the grandfather. He grinned”): 2 nouns of the same gender-bias.
 - Gender-Neutral (e.g., “The teacher billed the patient. He groaned”): 2 nouns equally gender-neutral.
- *Filler Items* (e.g., “The boy fed the teacher. He grinned”): 1 gender-biased noun and 1 gender-neutral noun. In this condition participants can use either gender information or verb-bias information to determine the pronoun’s referent. This condition was designed in order to encourage participants to consider both types of linguistic cues rather than always relying on either a gender cue or a verb-bias cue. We exclude these items from all analyses.

All of the sentences that included a male- or female-biased noun were counterbalanced for gender location (half male as subject, half female as subject). For example, in the Directly Determined condition we presented an equal number of female-male stimulus items (e.g., “woman”-“boy”) and male-female stimulus items (e.g., “boy”-“woman”).

We additionally generated 200 sentences that included a pronoun and a past-tense verb (e.g., “He grinned”). Half of the sentences include the male pronoun “he” and half the female pronoun “she”. These sentences were paired with the sentences described above to form the stimuli illustrated above and presented to subjects.

Since the focus of the experiment is on the influences of gender cue and verb-bias information on the resolution of a pronoun’s referent, we controlled verb-bias in the sentences across each condition. To determine the verb-bias within each sentence pair, we conducted a pretest survey (n=18 healthy young adults) by replacing the nouns of each sentence with “X” and “Y” (e.g., “The X hugged the Y. She smiled.”), and we asked participants to decide whether the pronoun refers to character “X” or “Y”. On average, the object noun of each sentence was preferred (70%) and this was consistent across each experimental condition (range 68.2% – 71.8%). This is consistent with the classes of verbs used in this study. The verb-bias ratings for each experimental condition were entered into a paired-samples t-test between each pair of experimental conditions and this analysis confirmed that the object noun was equally preferred across all experimental conditions (all tests $p > 0.1$).

A potential consequence of the object-preferred bias is that in half of the Determined items the correct referent is in the subject position and therefore these sentences may yield a conflict between the noun with the correct gender and the object-noun that is preferred by the verb. We consider this potential conflict in both our behavioral and fMRI analyses below, and we demonstrate that this has no consequence on behavioral performance or neural activation.

Lastly, to minimize any bias for selecting one noun over another noun, we statistically confirmed that both nouns within each experimental sentence were balanced for familiarity [$t(159)=1.05$], imageability [$t(159)=0.73$], word length [$t(159)=0.86$], and written lexical frequency [$t(159)=0.50$]. We also controlled for semantic-relatedness of the noun pairs across each condition. Semantic-relatedness between each noun pair (e.g., “boy”-“grandfather”) was quantified using the WordNet::Similarity package (Pedersen, Patwardhan, & Michelizzi, 2004), which generates a vector between two words on a scale of 0–1 reflecting semantic-relatedness, and has been demonstrated to correlate strongly with

human judgment ratings of semantic-relatedness (Patwardhan & Pedersen, 2006). A one-way ANOVA of Semantic-Relatedness with Experimental Condition as a factor revealed that Semantic-Relatedness was consistent across all conditions [$F(3,156)=1.89$, $p>0.1$].

Behavioral Procedure

Each experimental trial was comprised of a blank screen (2500 ms), a fixation cross (500 ms) and three stimulus events, each lasting the duration of one TR (3000 ms). In the first stimulus event, we presented a simple transitive sentence containing two nouns (e.g., “The woman paid the boy”). In the second stimulus event, we added a simple intransitive sentence containing a pronoun (“he” or “she”) and a past-tense verb (e.g., “He pouted”). In the final stimulus event, we added a probe with a question mark and asked participants to choose whether the pronoun refers to the left noun (e.g., “woman”) or the right noun (e.g., “boy”). Participants were given up to 5500 ms to respond (comprised of the concatenation of the final 3000 ms stimulus presentation and the following 2500 ms blank screen that preceded the fixation cross of the next event). To minimize task-related working memory demands, the linguistic materials presented in each stimulus event remained on the display screen for the full duration of the experimental trial. Thus, when participants made a response they had visual access to the transitive sentence containing both nouns, the sentence containing the pronoun, and the two noun choices.

Prior to the MRI scanning session, participants were given a practice session containing 10 trials to familiarize them with the form of the stimulus materials and the task, and we allowed them to ask questions before entering the scanner. These stimulus items were not re-presented in the experimental task.

All 200 stimulus materials (40 Directly Determined, 40 Indirectly Determined, 40 Undetermined (Gender-Biased), 40 Undetermined (Gender-Neutral) and 40 Fillers) were pseudo-randomly distributed across 5 experimental blocks of equal duration (approximately 9 minutes) comprised of 40 experimental stimuli each. There were an equal number of stimuli from each experimental condition (8 stimuli from each of 5 conditions) distributed over each experimental block. Items were presented using an event-related design that included null events for 15% of trials of varying durations (3000, 6000, and 9000 ms). Null events consisted of a blank (white) screen presentation. Each experimental block was followed by a 2 minute break.

Sentences were presented in black font on a white background using a mirror projection system connected to the computer running E-Prime presentation software. Using a fiber optic response pad (FORP), we monitored whether participants selected the subject or object noun. The FORP rested on the participant’s lap and contained 4 buttons oriented in a left-right fashion. Participants pushed the leftmost button with their left forefinger for a subject-noun response and the rightmost button with their right forefinger for an object-noun response. Since we only used sentence materials in the active voice, the subject-noun always appeared on the left side of the visual display and the object-noun always appeared on the right side of the visual display.

MRI Acquisition & Analysis

Scans were acquired on a Siemens 3.0T Trio scanner. Each session began with acquisition of a high-resolution T1-weighted structural volume using an MPRAGE protocol (TR = 3000 ms, TE = 3 ms, flip angle = 15°, 1 mm slice thickness, 192 × 256 matrix, resolution = .9766 × .9766 × 1 mm). A total of 865 BOLD fMRI whole brain volumes were acquired, with each volume containing 42 axial slices and acquired with fat saturation, 3 mm isotropic voxels, flip angle of 15°, TR = 3 s, TE_{eff} = 30 ms, and a 64 × 64 matrix.

BOLD fMRI data preprocessing and statistical analyses were performed using SPM5 (Wellcome Trust Centre for Functional Neuroimaging, London, UK). We first modeled each individual participant's data. Low-frequency drifts were removed with high-pass filtering with a cutoff period of 128 seconds and autocorrelations modeled using a first-order autoregressive model. Whole brain volumes for each participant were realigned to the first volume in the series (Friston et al., 1995) and coregistered with the structural volume (Ashburner & Friston, 1997). After realignment, we inspected each participant's motion in all directions and excluded two participants who had excessive motion artefact, defined as more than 3mm movement in any axis. The transformation required to bring a participant's images into standard MNI152 space was calculated using tissue probability maps (Ashburner & Friston, 1997), and these warping parameters were then applied to all functional brain volumes for that participant. During spatial normalization, functional data were interpolated to isotropic 2 mm voxels. The data were spatially smoothed with an 8 mm FWHM isotropic Gaussian kernel.

For each stimulus category, hemodynamic response was estimated by convolving the onset times with a canonical hemodynamic response function. A general linear model approach was used to calculate parameter estimates for each variable for each subject, and linear contrasts for comparisons of interest. These estimates were then entered into second-level random effects analyses to allow us to make inferences across participants.

In our initial analyses, we report contrasts using the Directly Determined condition as a high-level baseline since pronoun reference in this condition is unambiguous, and stimulus materials very closely matched to the items of experimental interest allowed us to make reasonably specific inferences about the basis for the observed differences in activation. In an initial comparison, we contrasted Determined (Directly and Indirectly) sentences with a correct referent in the subject position relative to sentences with the correct referent in the object-position. This contrast revealed only that all Determined stimuli in the subject position yield greater activation in right primary motor cortex (peak coordinates: 36, -13, 54; $z=6.08$; $p<0.001$ FDR-corrected), consistent with the button-press motor response. A similar pattern was observed for more restricted analyses of only Directly Determined stimuli (peak coordinates: 40, -14, 62; $z=5.52$; $p=0.002$ FDR-corrected) and only Indirectly Determined stimuli (peak coordinates: 38, -20, 38; $z=6.76$; $p<0.001$ FDR-corrected). These observations suggest that the same neural mechanisms, with the exception of the motor response required to make a button press, are recruited when gender information about nouns can be used to correctly identify a pronoun's referent in the subject and object positions of a sentence. Since we did not observe differences between subject and object referents in these sentences beyond the motor response, we collapsed across noun position for all additional contrasts. For all of these contrasts, we report regions of activation which exceed a $p<0.01$ FDR-corrected threshold and a 20 voxel extent to reduce the likelihood of the observation of false positive activation patterns. In some hypothesis-driven and theoretically-motivated contrasts with closely-matched baselines that differ from the Directly Determined baseline described above, we use a more liberal threshold ($p<0.001$ uncorrected) and only report activation of clusters which contained a peak voxel that exceeds a threshold of $z=3.69$ ($p<0.0001$ uncorrected). Additionally, we describe the theoretical motivation for using these different baselines in more detail in the Results Section below.

Results & Discussion

Behavioral Results

We evaluated response accuracy in the Directly and Indirectly Determined conditions since these conditions had a pronoun with a specifiable referent. Performance was highly accurate

for both types of Determined sentences (Directly: $M=99.28\%$, $SE=0.00$; Indirectly: $M=97.84\%$, $SE=0.01$). Additionally, accuracy was also high when the correct referent was in the Subject (97.8% , $SE=0.01$) and the Object ($M=99.3\%$, $SE=0.00$) position. These findings emphasize the usefulness of gender information when determining a correct referent even when the noun with corresponding gender is in the subject position and the verb is biased toward the noun in the object position.

To evaluate the relative uses of semantic gender cues and verb-bias information, we analyzed the proportion of participants' responses assigning the referent to the preferred object position (as demonstrated in the verb-bias pretest norms) in the Undetermined stimuli containing two gender-neutral nouns (e.g., proportion of responses of "patient" in the stimulus "The teacher billed the patient. He groaned") or containing two gender-biased nouns (e.g., "...boy...grandfather." "He..."). As predicted, when gender cues were not informative, participants preferred the object position in both types of Undetermined stimuli (gender-neutral: $M=72.6\%$, $SE=0.028$; gender-biased: $M=67.1\%$, $SE=0.041$). These findings for Undetermined stimuli are consistent with our pretest norms and suggest that in the absence of gender information participants tend to use verb-bias information in order to determine a pronoun's referent.

In summary, these behavioral findings suggest that individuals use gender information to select a referent of a pronoun when the gender information is informative, and they are highly accurate in their performance. When gender information is not informative, however, participants tend to rely on verb-bias information to inform their decision about a pronoun's referent.

BOLD fMRI Results

To determine the neural mechanisms that support the process of assigning a pronoun to an undetermined referent, we first compared all of the Undetermined sentences in which gender information is not informative (noun pairs that are both gender-biased or both gender-neutral) relative to the Directly Determined stimuli in which lexical semantic information can be used. This contrast, which emphasizes resources required to compute verb-bias information, revealed activation of both linguistic and decision-making networks during pronoun referent resolution (Table 1; Figure 1.A). Thus, we observed activation in left inferior frontal and left middle temporal cortices. These regions are consistent with a core language processing network, where left middle temporal cortex may contribute to interpreting the meaning of the verb in the sentence, and left inferior frontal cortex may play a role in the long-distance syntactic relationships relating the verb to a noun in the sentence (Kaan & Swaab, 2002). Activations consistent with a decision-making component included dlPFC, IPC and right inferior frontal cortex. Left dlPFC activation was predicted in our strategic account to support a probabilistic mechanism that helps determine the likelihood of a pronoun referring to one of two nouns and dlPFC has often been implicated in the evaluation of probability (Casey et al., 2001; Scheibe et al., 2010). We additionally observed right IPC activation, a brain region that has been associated with integrating probabilistic assessment with other aspects of the decision-making material (Huettel et al., 2005). Right inferior frontal recruitment also was observed, and although not predicted, this region may be related to increased processing demands for more difficult sentences (Just et al., 1996).

Since dlPFC has been implicated in explicit decision-making, we also performed this contrast for the preceding passive reading event. When participants were passively reading during these conditions, we observed a similar pattern of activation (Table 1, Figure 1.B). Thus, as described above, we found activation of a core language processing network, including left inferior frontal and left middle temporal cortices. We also found activations consistent with a decision-making network, including left dlPFC to support a probabilistic

mechanism (Casey et al., 2001; Scheibe et al., 2010) concerned with helping to determine the likelihood that a pronoun refers to one of the two nouns, right temporal-parietal-occipital activation consistent with an integrating mechanism (Dorris & Glimcher, 2004; Glimcher et al., 2005), and right inferior frontal activation that may play a role in working memory during language processing (Cooke et al., 2006). Since we observe similar activation patterns during the explicit decision and passive reading events, it is unlikely that the recruited regions can be explained entirely by explicit decision-making during task performance.

We also observed activation of both core language and decision-making regions in contrasts of each Undetermined condition with the Directly Determined baseline (Table 1; Figure 1.C and 2.D). For both the gender-neutral and gender-biased conditions, we observed activations consistent with a core language processing network. This included left inferior frontal recruitment which may support processing the syntactic relationships between a verb and a noun (Kuperberg, Lakshmanan, Caplan, & Holcomb, 2006a), and left middle temporal regions which may support processing lexical semantic (gender) information (Miceli et al., 2002). We also observed recruitment of regions implicated in a decision-making network. This included left dlPFC and left OFC, which we hypothesize may respectively contribute to the probabilistic assessment of the relationship between a pronoun and a pair of equally likely referents, and value assessment related to the risk of choosing the incorrect referent when a gender is not informative. Bilateral dorsal inferior frontal cortex may support additional working memory demands associated with computing the risk-related alternatives associated with a potentially ambiguous pronoun. This activation may be related more generally with the increased processing demands to comprehend more difficult sentences. This was seen in a previous study of sentence comprehension, where Cooke et al. (2006) observed increased recruitment of dorsal inferior frontal cortex related to increased grammatical processing demands. The gender-neutral condition additionally revealed left temporal activation which may support increased semantic demands associated with retrieving gender information about gender-neutral nouns, and occipital activation which may be associated with up-regulation of the entire sentence reading network to support the increased processing demands of this condition relative to the directly determined baseline. These observations again suggest that a large-scale neural network involving both language processing and decision-making mechanisms is recruited to support pronoun reference resolution.

In order to highlight specific aspects of our hypothesized large-scale neural network for resolving a pronoun's referent, we performed three additional contrasts. First, we directly compared Undetermined stimuli containing two nouns with the same gender bias (e.g., "boy" and "grandfather") to sentences containing two gender-neutral nouns (e.g., "teacher" and "patient"; see Table 2, Figure 2.A). This contrast revealed significantly greater activation in left insula extending to inferior frontal cortex and activation of dorsal inferior and precuneus regions. This pattern of activation is consistent with a verb-bias information account for determining the referent when gender-information is available (e.g., it is known that "boy" and "grandfather" are both male) but not informative for selecting between two possible referents (i.e., both are male). Specifically, the reader must infer the referent using information about the verb's subject or object preference. Inference making during sentence processing has been attributed to insula and inferior frontal cortex activation when readers can more easily make inferences relative to when it is more difficult to make inferences (Kuperberg, Lakshmanan, Caplan, & Holcomb, 2006b).

Second, we performed the reverse contrast to compare Undetermined stimuli containing two gender-neutral nouns relative to sentences containing two nouns with the same gender bias (Table 2, Figure 2.B). This contrast revealed significantly greater activation in bilateral

dIPFC. This is consistent with a probability mechanism which has been reported previously in non-language tasks to be supported by dIPFC (Casey et al., 2001; Scheibe et al., 2010). Specifically, the nouns in the gender-neutral condition by design do not have a strong gender bias cue and therefore individuals must calculate the probability for each noun (e.g., “patient”) being a male or female.

Third, to evaluate the role of the predicted risk mechanism that may support pronoun resolution, we contrasted the Indirectly Determined with the Directly Determined condition. Unlike the Undetermined conditions when the pronoun can correctly refer to either noun, the pronoun in the Indirectly Determined condition can refer correctly only to one noun. Specifically, the one gender-biased noun in the sentence “The mom served the child” has a conflicting gender value relative to the pronoun in the subsequent sentence “He pouted.” Therefore, we argue that there is some risk associated with choosing “child” as the referent of “she” here since this gender-neutral noun could be either a female or male. The imaging analysis (Figure 2.C, Table 2) revealed a cluster in the left hemisphere with a peak in ventrolateral prefrontal cortex (BA 47) that extends to include left OFC (BA 11; $x = -40$, $y = 34$, $z = -17$). OFC activation for Indirectly Determined stimuli, consistent with our prediction associating OFC with risk-evaluation. OFC is regularly implicated in non-linguistic neuroimaging studies involving risk (Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Fiddick, Spampinato, & Grafman, 2005; Hsu, Bhatt, Adolphs, Tranel, & Camerer, 2005; Sanfey, Loewenstein, McClure, & Cohen, 2006), and patients with atrophy or disease in OFC have deficits evaluating risk (Adolphs, 1999; Rogers et al., 1999; Rolls, 2000; Rosen et al., 2005; Salmon et al., 2003).

For this contrast of Indirectly Determined materials we also observed activation of right dIPFC, which may support additional recruitment of a probabilistic mechanism to support the calculation of the likelihood that the gender-neutral noun is a male or female. The activation of dIPFC, in the right hemisphere homologue, is consistent with the recruitment of this region in the two gender-neutral noun conditions when no gender-information is available.

Parametric Analysis

We observed in our behavioral analysis that individuals rely on gender cues regardless of verb-bias information. A further prediction that we can make from our model is that decision-making mechanisms should be up-regulated as gender cues become less informative. To evaluate this prediction, we conducted a parametric whole brain analysis of the data reported above. This was accomplished by using the gender rating norms as a metric of a noun’s uncertainty. We used the norms from our noun gender-rating pretest described in the Methods section. This was comprised of values from 1–3 representing gender uncertainty, where a lower number represents a strongly gender-biased noun (either male or female) and a higher number represents a gender-neutral noun. Since each sentence contained two nouns and there is also verb-bias for the object position we generated a formula that reflects the overall uncertainty of the sentence. This includes the sum of the gender ratings for each noun and an additional weight to reflect the verb-bias for the object position, as demonstrated in the behavioral analysis for each type of sentence. The verb-bias weight in the formula was assigned a value of +0.5 if the preferred noun was in the object position, a value of –0.5 if it was in the subject position, and a value of 0.0 if there was not a preferred noun. Thus, we can represent uncertainty using the following formula in which a higher number reflects greater uncertainty: $Uncertainty = Rating_{Noun1} + Rating_{Noun2} + W_{Position}$

When assigning a parametric value of uncertainty to each scan based on the corresponding stimulus sentence, we observed that increasing activation scaled with increasing uncertainty

in regions predicted to contribute to decision-making, including bilateral dlPFC and OFC. This is consistent with previous neuroimaging studies on decision-making which have demonstrated that activation in dlPFC scales in magnitude with probabilistic likelihood (Scheibe et al., 2010) and that activation of OFC has been implicated in evaluating risk (Hsu et al., 2005). See Figure 3 and Table 3 for a summary of results. Lastly, we observed an increase in activation in bilateral inferior frontal and left superior temporal cortices, which may additionally reflect increased processing demands for the core language network (Kaan & Swaab, 2002). Importantly, this parametric analysis provides robust ($p < 0.01$ FDR-corrected) converging evidence for the recruitment of hypothesized decision-making mechanisms observed in our categorical contrasts.

General Discussion

When a pronoun has an undetermined referent, we observed activation of cortical regions commonly associated with both language processing and decision-making. We hypothesize that core language processing mechanisms alone are not sufficient for resolution of pronoun reference, and we argue that strategic decision-making resources also must be recruited to support language processing. Specifically, we discuss the role of inferior frontal and temporal cortex in a core language processing network, and dlPFC, OFC, and IPC in a decision-making network that computes the EU of linguistic choice.

Core language processing mechanisms for resolving a referent

The observation of inferior frontal cortex activation in each condition relative to the Directly Determined condition suggests that this region plays an integral role in resolving a pronoun's referent. Our observed activation of inferior frontal and insula regions to support increased structural demands during sentence processing is consistent with several studies demonstrating increased inferior frontal activation while reading syntactically complex sentences (Caplan et al., 2000; Constable et al., 2004; Cooke et al., 2002; Just et al., 1996). Inferior frontal cortex thus may help determine the long-distance, syntactically-mediated relationship between a verb and a noun in a sentence (Grodzinsky & Friederici, 2006; Grodzinsky & Santi, 2008).

It is possible that when readers use verb-bias information to determine a pronoun's referent they use "implicit causality". This refers to the process of inferring a causal link within a discourse and has been demonstrated to contribute to pronoun resolution (Garnham, Oakhill, & Cruttenden, 1992b; Garnham, Traxler, Oakhill, & Gernsbacher, 1996a; McKoon, Greene, & Ratcliff, 1993). Our observation of insula, dorsal inferior frontal, and precuneus recruitment for referential processing was also described in reports on using causal information during language processing (Ferstl & von Cramon, 2001; Kuperberg et al., 2006a) and work demonstrating the additional role of the left insula in ordered information in language and non-language domains (Iijima, Fukui, & Sakai, 2009; Jankowski, Scheef, Huppe, & Boecker, 2009; Riecker, Brendel, Ziegler, Erb, & Ackermann, 2008).

In the current study we also observed activation in middle and superior temporal regions which have previously been reported to contribute to visual processing demands associated with reading linguistic material (Howard et al., 1992; Price et al., 1994). However, given our high-level baseline of contrasting sentences that only differ by the gender of the noun, it is unlikely that our observed temporal cortex activation is simply a result of reading-related processing demands. Instead, we argue that the observed pattern of temporal cortex activation may support the lexical semantic demands associated with retrieving gender-based lexical semantic information for nouns. Previous investigations thus have reported activation of middle temporal cortex when evaluating the gender of a noun (Heim, 2008; Heim, Opitz, & Friederici, 2002; Miceli et al., 2002). Several investigations also have

implicated superior temporal cortex as contributing to the integration of lexical semantic and syntactic information during sentence comprehension (Friederici, Makuuchi, & Bahlmann, 2009; Grodzinsky & Friederici, 2006). In this case, identification of the gender of a noun referentially linked to a pronoun relies on a semantic query of the gender value of the noun as well as integrating this lexical semantic information with syntactic information. Our finding of the recruitment of these regions while resolving a pronoun's referent is consistent with the role of middle and superior temporal cortices in processing lexical semantic information during sentence processing.

Decision-making mechanisms support resolving a referent

In this study, we were particularly interested in investigating the contribution of brain regions not traditionally associated with language. We focused on a network of cortical regions that is hypothesized to contribute to decision-making, including dlPFC, OFC, and IPC, and sought to extend this decision-making component to language. This was observed in previous studies of ambiguous pronoun use (Hammer et al., 2007; Nieuwland et al., 2007), although the basis for these activations was not clear. One potential account for our observation of recruitment of neuroanatomic regions associated with decision-making may be related to the task demands involved in the forced-choice selection of one noun over another noun. Future work is required to evaluate our hypothesized model in a more naturalistic setting or a fully passive reading task to demonstrate that decision-making mechanisms contribute to naturalistic language processing. However, despite this caveat, there are several reasons to believe that the observed patterns of activation that support the decision-making process are distinct from narrow, task-related demands.

First, we observed recruitment of similar neuroanatomic regions in both the passive reading and explicit decision events. This suggests that the observed regions are recruited independently of the explicit decision.

Second, the cortical regions that are thought to contribute to decision-making activations during pronoun resolution in this study are neuroanatomically distinct from regions commonly reported for sentence comprehension tasks with an explicit decision compared to passive sentence comprehension tasks. For example, one study involving auditory sentence comprehension reported increased activation associated with active listening (involving an explicit decision) compared to passive listening (no decision) in medial superior temporal cortex, cingulate gyrus, and right insula (Hasson, Nusbaum, & Small, 2006). In another study investigating explicit (semantic judgment) versus implicit (lexical decision) processing during a semantic task, the only regions up-regulated to support explicit task demands were in bilateral superior temporal cortex (Ruff, Blumstein, Myers, & Hutchison, 2008). Together, the activation patterns observed in these studies do not overlap with the regions observed in the current study that we propose are related to decision-making. Importantly, several of the regions activated in the current experiment have also been reported during passive reading of sentences with referential ambiguity (Hammer et al., 2007).

Third, pairs of contrasts showed selective activations despite involving exactly the same methods in both the experimental condition and the baseline condition. Thus, task-related resources involved in explicit decision-making such as maintaining and orienting attention to a reading task and making a dichotomous decision about a button response were present in both the experimental and baseline conditions during the contrast of the Indirectly Determined and Directly Determined conditions that revealed OFC activation. By comparison, dlPFC activation was observed during the contrasts of stimuli with pairs of Undetermined nouns that are gender neutral compared to stimuli with Undetermined nouns with the same gender bias. These experimental and baseline conditions involved the same methods, and moreover, used the identical methods in both experimental and baseline

conditions as the previous contrast. It is difficult to attribute these hypothesized differences to task-related demands.

Fourth, we obtained converging evidence of recruitment of the same decision-making regions in the parametric analysis. This analysis factors out activation patterns associated with task-related performance. This is because all task-related resource demands were held constant across all of the experimental conditions that contributed to the analysis, yet we observed variation in regional brain activation depending on the experimental conditions.

Fifth, our statistical comparisons of activation across conditions in the decision-making event revealed selective recruitment of hypothesized neuroanatomic regions associated with decision-making. These comparisons were designed to test specific predictions about the hypothesized role of probability and risk and involved the subtraction of stimuli that were closely matched for linguistic factors (frequency, imageability, familiarity) and often only differed by one or two letters. For example, in our contrast designed to evaluate “risk” we compared activation of Indirectly Determined stimuli (e.g., “The mom served the child. He pouted”) minus Directly Determined stimuli (e.g., “The mom served the child. She pouted”) and in this example the stimulus materials only differ by a single letter: “he” versus “she”. Therefore we argue that neuroanatomic mechanisms that may be associated with making an explicit decision would unlikely differ across such carefully controlled stimuli and also unlikely to occur in predicted regions.

Together, given the extant evidence for active versus passive language processing observed in two events, distinct activation patterns in pairs of contrasts that involve the identical methods, the parametric analysis that identifies predicted changes that correlate with stimulus properties despite identical methods, the selective recruitment of neuroanatomical regions under conditions with tightly controlled stimulus materials, and the literature suggesting that the neuroanatomic regions recruited in this study are distinct from those reported to contribute to explicit decision-making during language processing, it is unlikely that the cortical regions recruited in the current study are merely a result of task-related resource demands.

Dorsolateral prefrontal cortex supports a probabilistic mechanism during language processing

Consider first the mechanistic role of dlPFC during language processing. We propose a novel interpretation of dlPFC activation in language by attributing this to a probabilistic mechanism. Specifically, we observe dlPFC activation during the resolution of pronoun reference in Undetermined sentences. Activation of this region thus is greatest when individuals are required to evaluate the likelihood of a pronoun referring to one of two gender-neutral nouns. Nouns like “child” and “patient” do not have an obvious gender bias, and when a pronoun must refer to a noun of this sort, we argue that the listener computes the probability that these nouns are associated with a male or a female. This pattern of activation is consistent with non-linguistic studies demonstrating increased activation of dlPFC with increased probabilistic demands while making a decision (Casey et al., 2001; Krain, Wilson, Arbuckle, Castellanos, & Milham, 2006; Scheibe et al., 2010). For example, Scheibe et al (2010) instructed participants to decide whether the larger of two numbers was on the left or the right side of a visual display. Their task consisted of two events, a single number (e.g., “3”) presentation followed by the presentation of two numbers (e.g., “3” and “8”), and the value of the single number in the first event had a prior probability of 50%, 75%, or 100% of predicting where the larger number would appear. They observed that as the prior probability of predicted the location of the larger number increased, dlPFC activation also increased parametrically. Importantly, Scheibe et al’s probabilistic manipulation was conducted implicitly rather than explicitly, suggesting that dlPFC’s involvement in

probabilistic decision-making is an implicit response that does not require explicit or conscious knowledge about probabilistic priors. Therefore it is unlikely that dlPFC activation observed in our study is related to explicit knowledge about assigning a pronoun to its referent.

This is also consistent with the observation of dlPFC activation during the probabilistic calculation of a syntactic structure when confronted with a temporary structural ambiguity (Novais-Santos et al., 2007). This task also involved passive reading without an explicit decision. Importantly, the interpretation of the role of dlPFC as a probabilistic calculator does not refute theories suggesting that this region is involved in response selection or strategic manipulation, but rather complements these approaches. This is related to the observation that dlPFC mediates response selection during semantic retrieval (Badre, Poldrack, Pare-Blagoev, Insler, & Wagner, 2005; Badre & Wagner, 2002). Thus, dlPFC may support probabilistic computations of an outcome during response selection and strategic processes associated with language processing.

dlPFC activation has also commonly been reported to support executive resources required to support working memory and task difficulty (Braver et al., 1997; Cohen et al., 1997). However, all comparisons in the present study involved short visually-presented sentences that required minimal working memory and the stimuli remained visible during each trial, so it is unlikely that a working memory mechanism can account for our observed pattern of dlPFC activation.

While we do not explicitly test the relative contributions of left dlPFC and right dlPFC, our observation of only right hemisphere dlPFC activation in the Indirectly Determined materials may be related to the up-regulation of resources to support a probabilistic mechanism under conditions of increased processing demands. In this condition, readers must evaluate the probability of a gender-neutral noun when there is some risk associated with choosing that noun as a referent. We may not have observed left dlPFC activation in this condition because it may be present to some extent to support probabilistic demands in all conditions and thus is not evident in a contrast when performing comparisons of closely matched linguistic materials. Another potential account for the right lateralization may be related to the interplay between probability and risk in the larger hypothesized decision-making network. Recent transcranial magnetic stimulation (TMS) studies have demonstrated that deactivation of right dlPFC yields greater risk-taking behavior (Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006) while artificial stimulation of right dlPFC yields reduced risk-taking behavior (Fecteau et al., 2007). In the current study, we observed right lateralized dlPFC when participants assigned a pronoun to a more risky referent – a pattern consistent with TMS evidence, while left dlPFC was recruited when risk did not impact the choice of a referent.

Orbital frontal cortex supports a risk-related value mechanism during language processing

We observed OFC activation when participants were required to select an uncertain referent. First, we observed OFC activation in both of the Undetermined conditions relative to the Directly Determined condition. This finding suggests that OFC is recruited to resolve pronoun reference under conditions of uncertainty. Second, we observed OFC recruitment when participants were forced to choose a gender-neutral noun in the Indirectly Determined sentences. In these sentences there is no uncertainty as to the referent – one of the antecedents is clearly wrong due to a gender mismatch. In a stimulus like “The mom served the child. He pouted.” for example, “he” does not refer to “mom.” Rather, uncertainty results from being required to choose a noun in which the gender is not known, even though the pronoun has a clear gender bias. We interpret our observation of OFC activation in

conditions involving risk as evidence that OFC is contributing to the calculation of risk associated with making an uncertain linguistic choice. This account extends the neuroimaging and patient literature to a linguistic domain, implicating OFC as a value-based or risk-based mechanism (Adolphs, 1999; Breiter et al., 2001; Fiddick et al., 2005; Hsu et al., 2005; Rogers et al., 1999; Rolls, 2000; Rosen et al., 2005; Salmon et al., 2003; Sanfey et al., 2006).

A variety of other mechanistic roles for OFC during decision-making have been proposed in the neuroimaging literature, including response selection (Elliott, Dolan, & Frith, 2000), task-switching (Braver, Reynolds, & Donaldson, 2003) and inhibitory control (Horn, Dolan, Elliott, Deakin, & Woodruff, 2003). It is possible that our observation of OFC activation during the selection of competing referents is at least partially related to demands associated with response selection, although a response selection account alone cannot explain why we observe selective OFC activation in the Undetermined and Indirectly Determined stimulus materials in comparison to a baseline that also requires response selection. There may be task-switching demands in the current experiment related to switching between using lexical semantic information about gender and syntactic information about positional-biases, but this account also cannot explain our observation of selective activation of OFC during the Indirectly Determined condition that involves only lexical semantic information. Lastly, our observation of OFC activation during the selection of a gender-neutral noun when a gender-biased noun is incorrect (e.g. “The mom served the child. He pouted.”) could be potentially related to switching from a gender-biased noun with the wrong gender value to a gender-neutral noun or to inhibition associated with avoiding the gender-biased noun. However, since we did not see activation of OFC in the Directly Determined sentences that also require inhibiting one of the referents (e.g., inhibiting a female noun when identifying the referent of the pronoun “he”), it is less likely that an inhibition mechanism can alone account for our observed pattern of activation.

The localization of our observed OFC (BA 11) activation patterns were often near or included activation in the neighboring ventrolateral prefrontal cortex (BA 47) region. Activation of OFC can be difficult to observe in BOLD fMRI due to potential susceptibility bias (Stenger, 2006) and these two regions can be difficult to discriminate due to high individual variability in within this portion of the frontal lobes (Chiavaras & Petrides, 2000). Nonetheless, despite these technical challenges, activation of both orbital and ventrolateral prefrontal cortex have been implicated in supporting a risk mechanism (Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Fiddick, Spampinato, & Grafman, 2005; Hewig et al., 2009; Hsu, Bhatt, Adolphs, Tranel, & Camerer, 2005; Sanfey, Loewenstein, McClure, & Cohen, 2006). Future research that investigates the specific function of each of these neighboring regions in the context of language processing will be a fruitful endeavor.

Inferior parietal cortex supports an integration mechanism during language processing

Our observation of IPC activation may be consistent with previous studies suggesting that IPC supports discourse processing. Almor, Smith, Bonilha, Fridriksson, & Rorden (2007) observed IPC activation when individuals encountered repeated names compared to pronouns in a task requiring discourse interpretation. These investigators argued that IPC supports increased integration costs associated with bringing together multiple representations throughout the discourse. In another neuroimaging study, Kuperberg et al (2006a) reported increased activation of IPC when processing semantically unrelated sentences which needed to be integrated into a discourse. A recent patient study provides converging evidence for the role of IPC in discourse. Corticobasal syndrome patients, who have IPC disease, were shown to have difficulty maintaining connectedness while narrating a wordless picture story, and the extent of disease in this parietal region was related to the

extent of their discourse limitations (Gross et al., 2010). Together, these studies suggest that IPC may play a role integrating information during discourse processing.

Our account for IPC activation observed in the current study also proposes that this region supports an integration mechanism. Specifically, we argue that IPC supports the integration of probabilistic and value information in the form of EU. Both human (Dorris & Glimcher, 2004; Glimcher et al., 2005) and non-human primate studies (Gold & Shadlen, 2001; Platt & Glimcher, 1999; Sugrue, Corrado, & Newsome, 2004) have demonstrated that activation in IPC scales with the expected outcome of a reward. For example, investigators demonstrated that the neural firing of cells in monkey intraparietal sulcus increased with an expected reward of juice, and that this signal was greatest during the expectation phase compared to the reward phase of each trial (Platt & Glimcher, 1999). While our task does not involve an explicit reward outcome, we argue that EU in the context of language processing reflects the combined likelihood of a particular interpretation and the value associated with communicating clearly by using the interpretation that is most likely to be correct. Future work may address the explicit involvement of EU during language processing in tasks that incorporate an explicit reward component in association with a linguistic choice.

Conclusions

In this paper we established neural evidence for a two-component model for resolving a pronoun's referent. The present study focuses on the recruitment of a decision-making mechanism by a core language processing device. This is one instance of a large-scale neural network for language processing that emphasizes the recruitment of additional resources by a core language processing mechanism, as needed, to support sentence interpretation. Other instances include resolving conflict between potential interpretations of a doubly-quantified sentence (McMillan et al., 2011), engaging working memory during the processing of sentences with long-distance syntactic relations (Cooke et al., 2006; Cooke et al., 2002), and planning resources during the resolution of a temporary structural ambiguity (Novais-Santos et al., 2007). While previous investigations have suggested that decision-making resources may be required for reference resolution (Hammer et al., 2007), the present study attempted to identify specific mechanistic roles for regions commonly implicated in decision-making. We argue that dlPFC contributes to the probabilistic demands associated with resolving an undetermined referent, that OFC contributes to evaluating the risk associated with choosing an incorrect referent and that IPC integrates probability and risk to generate the EU of a linguistic choice. Future research that addresses the contributions of decision-making mechanisms for language processing, and that address the interaction between core language processing mechanisms and decision-making mechanisms will elucidate our overall understanding of the neurobiology of language.

Highlights

The results of this study suggest that core language and strategic decision-making resources are recruited during the resolution of a pronoun's referent:

- Lexical semantic resources in superior temporal cortex are recruited to determine the gender of a potential referent.
- Grammatical resources in inferior frontal cortex are recruited to determine the grammatical position of a potential referent.
- A probabilistic mechanism in dorsolateral prefrontal cortex is recruited when readers must evaluate which noun is more likely to be a referent in the absence of lexical semantic gender information (e.g., "The teacher billed the patient. She groaned.")

- A risk mechanism in orbitofrontal cortex is recruited when readers must choose a referent that has unknown gender (e.g., “The mom served the child. He pouted.”)
- An integration mechanism in inferior parietal cortex is recruited to integrate probability and risk in order to choose a referent that maximizes the expected utility of a decision.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the following grants from the National Institutes of Health: HD060406, NS44266, AG17586, AG15116, AG32953, and NS53488.

References

- Adolphs R. Social cognition in the human brain. *Trends in cognitive sciences*. 1999; 3:469–479. [PubMed: 10562726]
- Almor A. Noun-phrase anaphors and focus: the informational load hypothesis. *Psychol Rev*. 1999; 106(4):748–765. [PubMed: 10560327]
- Almor A, Smith D, Bonilha L, Fridriksson J, Rorden C. What is in a name? Spatial brain circuits are used to track discourse references. *Neuroreport*. 2007; 18(12):1215–1219. [PubMed: 17632270]
- Ashburner J, Friston K. Multimodal image coregistration and partitioning--a unified framework. *NeuroImage*. 1997; 6(3):209–217. [PubMed: 9344825]
- Badre D, Poldrack R, Pare-Blagoev E, Insler R, Wagner A. Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*. 2005; 47(6):907–918. [PubMed: 16157284]
- Badre D, Wagner A. Semantic retrieval, mnemonic control, and prefrontal cortex. *Behav Cogn Neurosci Rev*. 2002; 1(3):206–218. [PubMed: 17715593]
- Barraclough D, Conroy M, Lee D. Prefrontal cortex and decision making in a mixed-strategy game. *Nat Neurosci*. 2004; 7(4):404–410. [PubMed: 15004564]
- Braver T, Cohen J, Nystrom L, Jonides J, Smith E, Noll D. A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage*. 1997; 5(1):49–62. [PubMed: 9038284]
- Braver T, Reynolds J, Donaldson D. Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron*. 2003; 39(4):713–726. [PubMed: 12925284]
- Breiter H, Aharon I, Kahneman D, Dale A, Shizgal P. Functional Imaging of Neural Responses to Expectancy and Experience of Monetary Gains and Losses. *Neuron*. 2001; 30:619–639. [PubMed: 11395019]
- Camerer C. Behavioural studies of strategic thinking in games. *Trends in cognitive sciences*. 2003
- Caplan D, Alpert N, Waters G, Olivieri A. Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Hum Brain Mapp*. 2000; 9(2):65–71. [PubMed: 10680763]
- Caramazza A, Grober E, Garvey C, Yates J. Comprehension of Anaphoric Pronouns. *Journal of Verbal Learning and Verbal Behavior*. 1977; 16(5):601–609.
- Casey B, Forman S, Franzen P, Berkowitz A, Braver T, Nystrom L, et al. Sensitivity of prefrontal cortex to changes in target probability: a functional MRI study. *Hum Brain Mapp*. 2001; 13(1):26–33. [PubMed: 11284044]
- Chiavaras MM, Petrides. Orbitofrontal sulci of the human and macaque monkey brain. *Journal of Comparative*. 2000; 422:35–54.
- Clark R. *Meaningful Games: Exploring Language with Game Theory*. 2011:1–425.

- Clark R, Parikh P. Game theory and discourse anaphora. *Journal of Logic, Language and Information*. 2007; 16:265–282.
- Clark R, Parikh P, Ryant N. *Evolving Linguistic Defaults*. 2007
- Cohen J, Perlstein W, Braver T, Nystrom L, Noll D, Jonides J, et al. Temporal dynamics of brain activation during a working memory task. *Nature*. 1997; 386(6625):604–608. [PubMed: 9121583]
- Constable R, Pugh K, Berroya E, Mencl W, Westerveld M, Ni W, et al. Sentence complexity and input modality effects in sentence comprehension: an fMRI study. *NeuroImage*. 2004; 22(1):11–21. [PubMed: 15109993]
- Cooke A, Grossman M, DeVita C, Gonzalez-Atavales J, Moore P, Chen W, et al. Large-scale neural network for sentence processing. *Brain Lang*. 2006; 96(1):14–36. [PubMed: 16168473]
- Cooke A, Zurif E, DeVita C, Alsop D, Koenig P, Detre J, et al. Neural basis for sentence comprehension: grammatical and short-term memory components. *Hum Brain Mapp*. 2002; 15(2): 80–94. [PubMed: 11835600]
- Crawley R, Stevenson R, Kleinman D. The Use of Heuristic Strategies in the Interpretation of Pronouns. *Journal of Psycholinguistic Research*. 1990; 19(4):245–264. [PubMed: 2231480]
- Dorris M, Glimcher P. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron*. 2004; 44(2):365–378. [PubMed: 15473973]
- Elliott R, Dolan R, Frith C. Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cereb Cortex*. 2000; 10(3):308–317. [PubMed: 10731225]
- Fecteau S, Knoch D, Fregni F, Sultani N, Boggio P, Pascual-Leone A. Diminishing risk-taking behavior by modulating activity in the prefrontal cortex: a direct current stimulation study. *Journal of Neuroscience*. 2007; 27(46):12500–12505. [PubMed: 18003828]
- Ferstl E, von Cramon D. The role of coherence and cohesion in text comprehension: an event-related fMRI study. *Brain Res Cogn Brain Res*. 2001; 11(3):325–340. [PubMed: 11339984]
- Fiddick L, Spampinato M, Grafman J. Social contracts and precautions activate different neurological systems: An fMRI investigation of deontic reasoning. *NeuroImage*. 2005; 28:778–786. [PubMed: 15994098]
- Friederici A, Makuuchi M, Bahlmann J. The role of the posterior superior temporal cortex in sentence comprehension. *Neuroreport*. 2009; 20(6):563–568. [PubMed: 19287322]
- Friston K, Ashburner J, Frith C, Poline J-B, Heather J, Frackowiak R. Spatial registration and normalization of images. *Human Brain Mapping*. 1995; 3(3):165–189.
- Garnham A. Models of processing: discourse. *WIREs Cogni Sci*. 2010; 1(6):845–853.
- Garnham A, Oakhill J, Cruttenden H. The Role of Implicit Causality and Gender Cue in the Interpretation of Pronouns. *Lang Cogn Process*. 1992a; 7(3–4):231–255.
- The Role of Implicit Causality and Gender Cue in the Interpretation of Pronouns. 7 Cong. Rec. 1992b: 231–255.
- The locus of implicit causality effects in comprehension. 35 Cong. Rec. 1996a:517–543.
- Garnham A, Traxler M, Oakhill J, Gernsbacher M. The locus of implicit causality effects in comprehension. *Journal of memory and language*. 1996b; 35:517–543.
- Garvey C, Caramazza A. Factors influencing assignment of pronoun antecedents* 1. *Cognition*. 1975
- Glimcher P, Dorris M, Bayer H. Physiological utility theory and the neuroeconomics of choice. *Games and Economic Behavior*. 2005; 52(2):213–256. [PubMed: 16845435]
- Gold J, Shadlen M. Neural computations that underlie decisions about sensory stimuli. *Trends Cogn Sci*. 2001; 5(1):10–16. [PubMed: 11164731]
- Grodzinsky Y, Friederici A. Neuroimaging of syntax and syntactic processing. *Curr Opin Neurobiol*. 2006; 16(2):240–246. [PubMed: 16563739]
- Grodzinsky Y, Santi A. The battle for Broca's region. *Trends Cogn Sci*. 2008; 12(12):474–480. [PubMed: 18930695]
- Gross R, Ash S, McMillan C, Gunawardena G, Anderson C, Libon D, et al. Impaired information integration contributes to communication difficulty in corticobasal syndrome. *Cognitive and Behavioral Neurology*. 2010; 23:1–7. [PubMed: 20299856]

- Hammer A, Goebel R, Schwarzbach J, Münte T, Jansma B. When sex meets syntactic gender on a neural basis during pronoun processing. *Brain Research*. 2007; 1146:185–198. [PubMed: 16904083]
- Hasson U, Nusbaum H, Small S. Repetition suppression for spoken sentences and the effect of task demands. *J Cogn Neurosci*. 2006; 18(12):2013–2029. [PubMed: 17129188]
- Heim S. Syntactic gender processing in the human brain: a review and a model. *Brain Lang*. 2008; 106(1):55–64. [PubMed: 18226832]
- Heim S, Opitz B, Friederici A. Broca's area in the human brain is involved in the selection of grammatical gender for language production: evidence from event-related functional magnetic resonance imaging. *Neuroscience Letters*. 2002; 328:101–104. [PubMed: 12133565]
- Hewig J, Straube T, Trippel RH, Kretschmer N, Hecht H, Coles MG, Miltner WH. Decision-making under risk: an fMRI study. *Journal of Cognitive Neuroscience*. 2009; 21:1642–1652. [PubMed: 18823238]
- Horn N, Dolan M, Elliott R, Deakin J, Woodruff P. Response inhibition and impulsivity: an fMRI study. *Neuropsychologia*. 2003; 41(14):1959–1966. [PubMed: 14572528]
- Howard D, Patterson K, Wise R, Brown W, Friston K, Weiller C, et al. The cortical localization of the lexicons. Positron emission tomography evidence. *Brain*. 1992; 115(Pt 6):1769–1782. [PubMed: 1486460]
- Hsu M, Bhatt M, Adolphs R, Tranel D, Camerer C. Neural systems responding to degrees of uncertainty in human decision-making. *Science*. 2005; 310:1680–1683. [PubMed: 16339445]
- Huettel SA, Song AW, McCarthy G. Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *J Neurosci*. 2005; 25(13):3304–3311. [PubMed: 15800185]
- Iijima K, Fukui N, Sakai K. The cortical dynamics in building syntactic structures of sentences: an MEG study in a minimal-pair paradigm. *NeuroImage*. 2009; 44(4):1387–1396. [PubMed: 19041723]
- Indefrey P, Levelt W. The spatial and temporal signatures of word production components. *Cognition*. 2004; 92(1F2):101–144. [PubMed: 15037128]
- Jankowski J, Scheef L, Huppe C, Boecker H. Distinct striatal regions for planning and executing novel and automated movement sequences. *NeuroImage*. 2009; 44(4):1369–1379. [PubMed: 19059350]
- Just M, Carpenter P, Keller T, Eddy W, Thulborn K. Brain activation modulated by sentence comprehension. *Science*. 1996; 274(5284):114–116. [PubMed: 8810246]
- Kaan E, Swaab T. The brain circuitry of syntactic comprehension. *Trends in cognitive sciences*. 2002; 6(8):350–356. [PubMed: 12140086]
- Kehler A, Kertz L, Rohde H, Elman J. Coherence and coreference revisited. *Journal of Semantics*. 2008; 25(1):1.
- Kennison SM, Trofe JL. Comprehending pronouns: a role for word-specific gender stereotype information. *J Psycholinguist Res*. 2003; 32(3):355–378. [PubMed: 12845944]
- Kiani R, Shadlen M. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*. 2009; 324(5928):759–764. [PubMed: 19423820]
- Kim J, Shadlen M. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat Neurosci*. 1999; 2(2):176–185. [PubMed: 10195203]
- Knoch D, Pascual-Leone A, Meyer K, Treyer V, Fehr E. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*. 2006; 314(5800):829–832. [PubMed: 17023614]
- Krain A, Wilson A, Arbuckle R, Castellanos F, Milham M. Distinct neural mechanisms of risk and ambiguity: a meta-analysis of decision-making. *NeuroImage*. 2006; 32(1):477–484. [PubMed: 16632383]
- Kuperberg G, Lakshmanan B, Caplan D, Holcomb P. Making sense of discourse: an fMRI study of causal inferencing across sentences. *NeuroImage*. 2006a; 33(1):343–361. [PubMed: 16876436]
- Making sense of discourse: an fMRI study of causal inferencing across sentences. *33 Cong. Rec*. 2006b:343–361.
- Leon M, Shadlen M. Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron*. 1999; 24(2):415–425. [PubMed: 10571234]

- Mayol L. Contrastive pronouns in null-subject Romance languages. *Lingua*. 2010; 120:2497–2514.
- Mayol L, Clark R. Pronouns in Catalan: Games of partial information and the use of linguistic resources. *Journal of Pragmatics*. 2010; 42:781–799. [PubMed: 20161616]
- McClure S, Laibson D, Loewenstein G, Cohen J. Separate neural systems value immediate and delayed monetary rewards. *Science*. 2004; 306(5695):503–507. [PubMed: 15486304]
- Discourse models, pronoun resolution, and the implicit causality of verbs. *19 Cong. Rec.* 1993:1040–1052.
- McMillan, CT.; Ryant, N.; Coleman, D.; Clark, R.; Grossman, M. Strategic resources support the interpretation of doubly-quantified sentences. In: Carlson, L.; Hoelscher, C.; Shipley, TF., editors. *Expanding the Space of Cognitive Science: Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*; Boston, MA: Cognitive Science Society; 2011. p. 2433–2438.
- Miceli G, Turriziani P, Caltagirone C, Capasso R, Tomaiuolo F, Caramazza A. The neural correlates of grammatical gender: an fMRI investigation. *J Cogn Neurosci*. 2002; 14(4):618–628. [PubMed: 12126502]
- Miller GF, Todd PM. Mate choice turns cognitive. *Trends Cogn Sci*. 1998; 2(5):190–198. [PubMed: 21227154]
- Nieuwland MS, Petersson KM, Van Berkum JJA. On sense and reference: examining the functional neuroanatomy of referential processing. *NeuroImage*. 2007; 37(3):993–1004. [PubMed: 17611124]
- Novais-Santos S, Gee J, Shah M, Troiani V, Work M, Grossman M. Resolving sentence ambiguity with planning and working memory resources: Evidence from fMRI. *NeuroImage*. 2007; 37(1):361–378. [PubMed: 17574445]
- Patokos T. The relevance of Nash equilibrium to psychiatric disorders. *Theoretical Medicine and Bioethics*. (in press).
- Patwardhan, S.; Pedersen, T. Using WordNet Based Context Vectors to Estimate the Semantic Relatedness of Concepts; Paper presented at the EACL 2006 Workshop Making Sense of Sense - Bringing Computational Linguistics and Psycholinguistics Together; Trento, Italy. 2006.
- Paulus M, Hozack N, Zauscher B, McDowell J, Frank L, Brown G, et al. Prefrontal, parietal, and temporal cortex networks underlie decision-making in the presence of uncertainty. *NeuroImage*. 2001; 13:91–100. [PubMed: 11133312]
- Pedersen, T.; Patwardhan, S.; Michelizzi, J. WordNet::Similarity - Measuring the Relatedness of Concepts; Paper presented at the Nineteenth National Conference on Artificial Intelligence (AAAI); San Jose, CA. 2004.
- Pickering MJ, Majid A. What are implicit causality and consequentiality? *Language and Cognitive Processes*. 2007; 22(5):780–788.
- Pinker S, Nowak M, Lee J. The logic of indirect speech. *Proceedings of the National Academy of Sciences*. 2008; 105(3):833.
- Platt M, Glimcher P. Neural correlates of decision variables in parietal cortex. *Nature*. 1999; 400:233–238. [PubMed: 10421364]
- Price C, Wise R, Watson J, Patterson K, Howard D, Frackowiak R. Brain activity during reading. The effects of exposure duration and task. *Brain*. 1994; 117(Pt 6):1255–1269. [PubMed: 7820564]
- Riecker A, Brendel B, Ziegler W, Erb M, Ackermann H. The influence of syllable onset complexity and syllable frequency on speech motor control. *Brain Lang*. 2008; 107(2):102–113. [PubMed: 18294683]
- Rogers R, Owen A, Middleton H, Williams E, Pickard J, Sahakian B. Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *Journal of Neuroscience*. 1999; 19:2029–2038.
- Rolls E. The orbitofrontal cortex and reward. *Cerebral Cortex*. 2000; 10:284–294. [PubMed: 10731223]
- Rosen H, Allison S, Schauer G, Gorno-Tempini M, Weiner M, Miller B. Neuroanatomical Correlates of Behavioural Disorders of Dementia. *Brain*. 2005; 128:2612–2625. [PubMed: 16195246]
- Rudolph U, Försterling F. The psychological causality implicit in verbs: A review. *Psychological Bulletin*. 1997; 121(2):192–218.

- Ruff I, Blumstein S, Myers E, Hutchison E. Recruitment of anterior and posterior structures in lexical-semantic processing: An fMRI study comparing implicit and explicit tasks. *Brain and language*. 2008; 105(1):41–49. [PubMed: 18279947]
- Sally D. Risky speech: behavioral game theory and pragmatics. *Journal of pragmatics*. 2003
- Salmon E, Garraux G, Delbeuck X, Collette F, Kalbe E, Zuendorf G, et al. Predominant Ventromedial Frontopolar Metabolic Impairment in Frontotemporal Dementia. *NeuroImage*. 2003; 20:435–440. [PubMed: 14527604]
- Sanfey AG, Loewenstein G, McClure SM, Cohen JD. Neuroeconomics: cross-currents in research on decision-making. *Trends Cogn Sci (Regul Ed)*. 2006; 10(3):108–116. [PubMed: 16469524]
- Sanfey AG, Rilling J, Aronson J, Nystrom L, Cohen J. The neural basis of economic decision-making in the Ultimatum Game. *Science*. 2003; 300(5626):1755–1758. [PubMed: 12805551]
- Sanford A. Towards a psychological model of written discourse comprehension. *Advances in Psychology*. 1982
- Schall J. Neural basis of deciding, choosing and acting. *Nat Rev Neurosci*. 2001; 2(1):33–42. [PubMed: 11253357]
- Scheibe C, Ullsperger M, Sommer W, Heekeren HR. Effects of parametrical and trial-to-trial variation in prior probability processing revealed by simultaneous electroencephalogram/functional magnetic resonance imaging. *J Neurosci*. 2010; 30(49):16709–16717. [PubMed: 21148010]
- Shadlen M, Newsome W. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol*. 2001; 86(4):1916–1936. [PubMed: 11600651]
- Stenger, VA. Technical considerations for BOLD fMRI of the orbitofrontal cortex. In: Zald, D.; Rauch, S., editors. *The orbitofrontal cortex*. Oxford: Oxford University Press; 2008. p. 423–446.
- Sugrue L, Corrado G, Newsome W. Matching behavior and the representation of value in the parietal cortex. *Science*. 2004; 304(5678):1782. [PubMed: 15205529]
- Tyler L, Marslen-Wilson WD, Randall B, Wright P, Devereux BJ, Zhuang J, et al. Left inferior frontal cortex and syntax: function, structure, and behaviour in patients with left hemisphere damage. *Brain*. 2011; 134(2):415–431. [PubMed: 21278407]
- Vickery T, Jiang Y. Inferior parietal lobule supports decision making under uncertainty in humans. *Cereb Cortex*. 2009; 19(4):916–925. [PubMed: 18728197]

Figure 1.
fMRI activation patterns illustrating recruitment of core language and decision-making regions during various Undetermined sentences relative to Directly Determined sentences
NOTE
1. (A) Undetermined > Directly Determined during explicit decision event, (B) Undetermined > Directly Determined during passive reading event, (C) Gender-Neutral > Directly Determined, and (D) Gender-Biased > Directly Determined.

Figure 2.

Regions of activation in comparisons of specific Undetermined types of sentences¹

NOTE

1. (A) Gender-Biased > Gender-Neutral, (B) Gender-Neutral > Gender-Biased and (C) Indirectly Determined > Directly Determined.

Figure 3.
Regions of increased activation from parametric analysis of pronoun reference uncertainty.

Table 1

Regions of activation for Undetermined > Directly Determined during explicit decision, Undetermined > Directly Determined for passive reading, Gender-Neutral > Directly Determined, and Gender-Biased > Directly Determined¹

Neuroanatomic Region (BA)	L/R	Coordinates	Voxels	Z-score
Undetermined > Directly Determined Explicit Decision				
Inferior frontal (45)	R	54 22 17	283	4.94
Inferior frontal (44/45), dorsolateral prefrontal (46)	L	-54 20 8	976	4.61
Superior/Middle temporal (39/22)	L	-44 -50 14	199	4.49
Inferior parietal (40)	R	54 -53 28	117	4.43
Undetermined > Directly Determined Passive Reading				
Dorsolateral prefrontal (46), inferior frontal (45)	L	-55 28 13	555	4.16
Dorsal inferior frontal (9)	L	-36 12 40	212	4.14
Middle temporal (21)	L	-48 -37 0	120	3.81
Gender-Neutral > Directly Determined				
Inferior frontal (45)	R	54 22 17	543	4.81
Inferior frontal (44/45), dorsolateral prefrontal (46) orbitofrontal (47)	L	-44 16 18	1811	4.69
Middle temporal (21)	L	-48 -44 7	115	4.38
Occipital (17)	L	-14 -95 1	134	4.35
Dorsomedial prefrontal (8)	-	-4 35 35	52	4.29
Superior temporal (39)	R	48 -48 21	70	4.26
Middle occipital (18)	R	22 -93 5	105	4.10
Anterior temporal (20/21)	L	-57 -11 -18	21	4.06
Orbitofrontal (47/11)	R	52 25 -6	42	4.00
Middle Occipital (19)	R	36 -91 8	20	3.90
Anterior cingulate (32)	-	-6 23 39	30	3.83
Parietotemporal junction (39)	L	-40 -59 23	38	3.80
Gender-Biased > Directly Determined				
Inferior frontal (45)	R	55 20 17	300	5.49
Inferior frontal (45), orbitofrontal (47/11)	L	-57 20 8	1225	4.84

Neuroanatomic Region (BA)	L/R	Coordinates	Voxels	Z-score
Dorsomedial prefrontal (8)	-	-2 39 37	112	4.67
Middle temporal (21)	L	-46 -46 8	156	4.59
Dorsolateral prefrontal (6)	L	-38 6 44	74	4.31
Dorsomedial prefrontal (8/6)	-	-10 16 51	26	4.12
Frontal pole (10)	R	52 -50 21	41	3.98

NOTE

/ All contrasts significant at $p < 0.001$ FDR-corrected except Undetermined > Directly Determined passive reading significant at $p < 0.001$ uncorrected.

Table 2

Regions of activation for Gender-Biased > Gender-Neutral, Gender-Neutral > Gender-Biased, and Indirectly Determined > Directly Determined sentences.[/]

Neuroanatomic Region (BA)	L/R	Coordinates	Voxels	Z-score
Gender-Biased > Gender-Neutral				
Insula (22/6), Inferior frontal (44)	L	-55 -1 13	622	4.74
Putamen	L	-20 2 4	34	4.37
Dorsal inferior frontal (6)	L	20 -13 56	37	4.13
Precuneus (7)	-	-10 -33 48	78	4.11
Anterior cingulate (24)	-	-14 34 11	154	4.07
Gender-Neutral > Gender-Biased				
Dorsolateral prefrontal (46)	L	-48 30 13	113	4.20
Inferior frontal (45), dorsolateral prefrontal (46)	R	50 22 21	74	3.98
Middle occipital (18)	R	16 -97 7	183	3.97
Middle occipital	L	-18 -95 5	93	3.62
Indirectly Determined > Directly Determined				
Orbitofrontal (47/11)	L	-51 33 -10	112	4.01
Middle occipital (18)	L	-30 -93 3	153	3.96
Inferior frontal (44), Dorsolateral prefrontal (46)	R	34 15 21	71	3.76

NOTE

[/] All activations significant at $p < 0.0001$ uncorrected.

Table 3

Regions of activation for parametric analysis of pronoun uncertainty^a

Neuroanatomic Region (BA)	L/R	Coordinates	Voxels	Z-score
Parametric Activation Associated with Increasing Uncertainty				
Inferior frontal (44/45)	R	54 20 16	550	4.86
Inferior frontal (44), orbitofrontal (11/47) dorsolateral prefrontal (46)	L	-44 16 18	1842	4.65
Dorsomedial prefrontal (8)	-	-2 39 37	140	4.44
Parietotemporal junction (39/22)	L	-48 -44 8	110	4.16
Parietotemporal junction (39/22)	R	-44 -50 16	34	4.01
Orbital frontal (47/11)	R	52 25 -6	61	4.11
Dorsomedial frontal (6)	-	-8 16 51	43	3.92
Anterior temporal (21)	R	-63 -9 -16	81	3.64
Superior temporal (22)	R	50 -35 5	33	3.36

NOTE

^aSignificant at p<0.01 FDR-corrected.