Noname manuscript No.
(will be inserted by the editor)

Focus Games

Jon Scott Stevens

the date of receipt and acceptance should be inserted later

Abstract This paper provides a game-theoretic analysis of contrastive focus, extending insights from recent work on the role of noisy communication in prosodic accent placement to account for focus within sentences, sub-sentential phrases (e.g. in "farmer sentences") and words. The shared insight behind these models is that languages with prosodic focus marking assign prosodic prominence only within elements which constitute material critical for successful interpretation. We first take care to distinguish the information-structural notion of (contrastive) focus from an ontologically distinct notion of givenness marking, and then outline the core properties of focus. We then introduce a signaling game between a speaker and hearer in which the goal is to transmit semantic content with the smallest signal possible. We apply the Iterated Best Response (IBR) method of Franke (2011) to find equilibrium strategies in this game, where a unique equilibrium strategy in this case picks out the "critical information" of an utterance, which by hypothesis constitutes its focus. We show that iterating this game at different syntactic levels of a sentence makes correct predictions about the role of contrast in determining stress within words and phrases, and can be extended to account for association and second occurrence focus effects.

1 Introduction

Despite the view that communication is an "ancillary property" of human language (Chomsky, 2011, p.275), it is undeniable that language is used, among other things, to communicate propositional thought from a speaker to a hearer. But it is an imperfect mode of transmission. There is no guarantee that the speaker's utterance will be correctly perceived and interpreted by the hearer. Echoing Shannon's (1948) classic work in information theory—more recently applied to pragmatics by Benz (2012)—linguistic communication, like any form of information transmission, is subject to *noise*, the random alteration or deletion of parts of a signal.

Interlocutors' knowledge that the information contained within their utterances may be altered by noise can place additional pragmatic constraints on how to opti-

Address(es) of author(s) should be given

mally formulate an utterance. One domain in which this is apparent is the study of information structure, loosely defined as those areas where discourse context partially determines linguistic form. This paper extends recent work (Schmitz, 2008; Bergen & Goodman, 2014) considering the role of noise in determining the placement of prosodic *focus*, conceived of as the contrastively interpreted assignment of pitch accents within a sentence. We use the idea of noisy communication to motivate a game-theoretic model of the selection of pitch accent assignment domains. The model derives key facts about focus in answers to wh-questions as well as in cases of intra-sentential contrast ("farmer sentences") and accent shift within individual lexical items. Moreover, the model explains some surprising observations about focus placement in the presence of multiple competing antecedents (3.3) and in second occurrence contexts (4.2), as well as observations about the kinds of non-literal pragmatic inferences that can arise from the manipulation of focus structure (Section 5).

The remainder of Section 1 outlines our background assumptions concerning the nature of focus; Section 2 introduces the game-theoretic framework upon which our model is built; Section 3 presents the model; Section 4 gives suggested extensions to account for focus association and second occurrence focus; Section 5 discusses higher-level pragmatic inferences which can arise from the use of focus; Section 6 concludes.

1.1 Background

One of the primary goals of any theory of information structure is to explain the observation that intonation is systematically constrained by meaning in many languages. Going back to Halliday (1967), this relationship between meaning and prosody has been taken to involve some notion of focus, where a constituent is focused when the speaker wishes to "mark out" that part of his utterance "as that which he wishes to be interpreted as informative." (Halliday, 1967, p.204). Since that time focus has taken many forms. A debate between Chomsky (1971) and Bolinger (1972) raised the question of whether default rules of intonation play a part in determining sentence-level prosody, or whether focus is entirely semantically motivated. Most accounts (e.g. Vallduví 1990, Rooth 1992, Roberts 1996, Krifka 2001, Selkirk 2007, and many others) assume that there is an abstract focus representation such that the answer in (1-a) (small caps denotes strong prosodic prominence in the form of a pitch accent) has the structure represented in (1-b).

(1) a. Q: What class does Sue teach on Friday afternoons?
A: Sue teaches PHONETICS on Friday afternoons.
b. Sue teaches [phonetics]$_{\mathrm{F}}$ on Friday afternoons.

The intuition behind this structure is that the question sets up a context where the answer is expected to take the form 'Sue teaches x on Friday afternoons', where 'x' is some course which Sue is qualified to teach. In this sense only the choice of value for 'x' is really informative. More recently, Halliday's original insight—that focus serves to mark out the informative content—has been cashed out by considering the role of noise in linguistic communication. Schmitz (2008) formalizes the concept of "i-critical" material, a natural extension of Roberts's (1996) idea that foci serve as minimal answers to a question under discussion (QUD). Under this approach, (1-b) is licensed by the fact that the QUD systematically constrains the set of possible responses. Assuming a Hamblin (1973) semantics for questions, answers are confined to those represented in (2) below.

(2) What class does Sue teach on Friday afternoons?
{'Sue teaches phonetics on Friday afternoons',
'Sue teaches phonology on Friday afternoons', $\cdots$}

Because of this restriction, any material outside of the DP *phonetics* in the answer given in (1-a) is recoverable on the assumption that a cooperative speaker will provide an answer from that set. Thus, if noise were to obscure the other portions of the speaker's utterance, the hearer would still be able to reckon that the object DP denotation 'phonetics' is part of the QUD answer, which would pick out a unique proposition, 'Sue teaches phonetics on Friday afternoons'. Prosodic prominence on the object helps to ensure, via additional phonetic prominence, that noise does not obscure that crucial material. By that token, the focus pattern is licensed for the same reason that an elliptical answer is also possible: The non-prominent material is informationally redundant to the extent that it simply isn't worth the effort associated with additional prominence.

(3) Q: What class does Sue teach on Friday afternoons?
A: Phonetics.

Bergen & Goodman (2014) use tools from game-theoretic pragmatics, building on Franke (2009), Benz (2012) and others, to make the role of noisy communication more explicit. Under their account, pragmatic reasoning is used to derive exhaustive interpretations for contrastively stressed elements, as in the following, where the answer is intended as an exhaustive list of movie-goers, implying that no other contextually relevant individuals went with Bob.

(4) Q: Who went to the movies?
A: BOB went to the movies.

Under their account, if the subject "Bob" were obscured by noise, then the hearer might mistakenly suppose that Alice went to the movies, and therefore, strong prominence on "Bob" is used as a strategy by the speaker to reduce the likelihood of this false belief arising due to noise. If the hearer knows that the speaker is using such a strategy, then the hearer will infer that the speaker is specifically trying to prevent her from forming the belief that Alice went to the movies. Therefore, that proposition, as well as any other analogous propositions, must be false, and so the supplied answer must be an exhaustive list of relevant individuals who went to the movies.

It should be noted that the story is a bit more complex: it is possible to counter the exhaustivity implication associated with strong prosodic prominence on "Bob" by using a so-called B-accent or rise-fall-rise accent, explicitly signaling discourse incompleteness or non-finality and suggesting that the speaker did not provide all potentially relevant information (Jackendoff 1972, Büring 2003, Lai 2012, Wagner 2012a). But as a base case, the noise-based account of (4) provides a promising extension of Schmitz's insight that prosody is constrained by notions of communicative strategy and efficiency.

There are two *prima facie* problems with this general approach, and both of them are ameliorated by acknowledging that there are independent, inviolable facts about a language's grammar which operate at cross purposes to these communicative considerations. First, it is not the case that everything in focus is equally prosodically prominent. For example, the observed contour in (5-a) below is thought to have an underlying abstract structure in which the VP *is teaching phonetics* is marked as 'focused', but within that VP only the object *phonetics* receives particularly strong stress. In fact, analyses based on so-called focus projection going back to Chomsky (1971) have claimed that right-edge VP focus is prosodically identical to sentence-wide focus, where in both cases the strongest stress falls at the right edge of the clause (the default pattern), as in the following.

(5) a. Q: What is Sue doing?
A: Sue [is teaching PHONETICS]$_{\mathrm{F}}$
b. Q: What event is happening right now?
A: [Sue is teaching PHONETICS]$_{\mathrm{F}}$

Empirical investigations of this claim have yielded mixed results, suggesting that in some cases there may be a detectable difference between the two structures (De Kuthy & Meurers 2012, De Kuthy & Stolterfoht 2014). Nonetheless, it is clear that independent phonological parameters, which incorporate notions such as eurhythmy (e.g. avoidance of stress clash, or avoidance of adjacent strong prominences) and prosodic phrasing (see Truckenbrodt 2007, for an overview), operate either on top of or prior to the semantic-pragmatic processes which constrain the selection of abstract focus representations. In other words, grammar mediates between discourse considerations and phonetic output, and when grammar and discourse considerations collide, grammar wins.

Second, not all languages mark focus phonologically. Focus and other similar information-structural notions are often encoded syntactically (though often these syntactic categories do not have identical semantic-pragmatic reflexes to what is traditionally called focus; see Zimmermann & Onea 2011), and in many languages prosody is assigned rigidly, with no possibility to move or add prominence to different elements in a sentence. This does not require us to claim that languages like English and German, which do indeed mark focus prosodically, operate under different communicative principles. Rather, we only need to acknowledge that many languages have phonological grammars which *a priori* prohibit prosody from being exploited as a marker of focus. And in these cases it may be that other resources, be they syntactic or morphological, can be employed instead. While we focus here on prosodic prominence, a similar analysis may be possible for syntactic or morphological focus marking strategies. Though these grammatical reflexes do not necessarily contribute to phonetic prominence, different conceptions of prominence are possible, and it could be that, for example, a special syntactic position for focus serves a similar discourse function to prosodic focus marking in that it signals to the hearer where he/she will find the most important information for purposes of interpretation.

In sum, a combination of grammatical and pragmatic factors creates a complex relationship between meaning and prosodic prominence, but one potential explanatory factor in this relationship arises from the simple fact that language

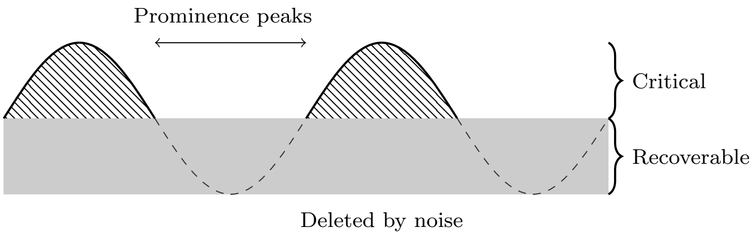

Fig. 1 An ideal signal gives more prominence to what is critical, and less prominence to what the hearer could recover from context.

is used to transmit information through a noisy channel. Shannon (1948, p.414) sums up how the effect of noise on communication can be ameliorated:

> An approximation to the ideal [signal transmission] would have the property that if the signal is altered in a reasonable way by the noise, the original can still be recovered. In other words the alteration will not in general bring it closer to another reasonable signal than the original. This is accomplished at the cost of a certain amount of redundancy in the coding.

In light of this, we may view prosodic accent as a form of redundancy, a tool to ameliorate the deleterious effect of the noisy channel.

This paper aims to further extend this insight by modeling the selection of prosodically accented material as part of a game between speaker and hearer, and aims to connect that account to well-established facts about focus placement. The *focus game* is a model of how to reason about what is "important" or "critical" for interpretation, with the linking hypothesis that this important material will be marked off as such via grammatical means (such as abstract phonological accent), and that this correlates with at least a partial increase in overall prominence, where prominent elements are assumed to be more obvious, and harder to obscure via noise. If the prominent material is the critical material for interpretation, and if the non-prominent material is recoverable by the hearer, then the signal reduces the effect of noise in a perfectly efficient way. Our view is that language tends to approach (but not meet) this ideal. Fig.1 illustrates.

Specifically, we explore the role of noise in focus placement beyond the simple question-answer pairs of (4) and (5). To this end, we move from a QUD-based notion of focus to what Selkirk (2007) and others call *contrastive focus*, which invokes a well-defined set of contrasting alternatives in the discourse context. This overlaps with the QUD-based conception of focus in that a QUD itself can be seen as a set of contrasting alternatives in cases where the set of possible answers is sufficiently restricted. This is the case for (1), where the context presumably provides a constrained set of courses that Sue could be teaching. Such cases are distinct from more open-ended question-answer contexts, e.g. 'What is something that Sue did today?', where the Hamblin set is almost completely open-ended, and thus one does not get the sense of selecting from among distinct alternatives. Selkirk (2007) and Katz & Selkirk (2011) suggest analyzing the assignment of pitch accent placement in these open-ended cases in terms of 'given' vs. 'new' rather than marking foci, giving evidence that the pitch prominences that do occur in these cases are lower in pitch, duration and intensity than contrastively interpreted foci.

Crucially, while some instances of contrastively interpreted focus are straightforwardly covered by an account along the lines of Schmitz (2008), a fuller account must be extended to cover cases like (6) and (7), where there is no obvious relationship between the QUD and the focusation of the answer.

(6) An AMERICAN farmer punched a CANADIAN farmer.

(7) We drove from MINNE-sota to SARA-sota.

In (6), we see contrastively interpreted foci on *American* and *Canadian* in what could be an out-of-the-blue utterance, or an answer to some broad, open-ended QUD like 'what happened?', which is expected to create default, non-contrastive prominence assignment throughout the whole clause (Selkirk 2007), indicated by unlabeled brackets in (8) below.[1] It is in virtue of the contrastive nature of the pair {'American farmer','Canadian farmer'} that the additional prominences are licensed. After Rooth (1992), we may analyze this in terms of contextually salient alternative NP denotations: The salient set of alternatives {'American farmer','Canadian farmer'} licenses focus marking within the NP domain on both of the adjectives, lending those elements higher relative prominence.

(8) [An [American]$_{\mathrm{F}}$ farmer punched a [Canadian]$_{\mathrm{F}}$ farmer]

Rooth provides a formal description of focus licensing in the form of alternative semantics: an "F-feature" is licensed on each adjective within the NP domain because each NP has an antecedent (the other NP) which is a member of the alternative set {'P farmer'}, obtained by substituting an open variable of the appropriate type for the denotation of the F-marked element. Under this analysis, F-features presuppose that such an alternative set is evoked, either explicitly or implicitly, by the discourse context. In (7), we have a similar case of contrastive focus, but one where the relevant focus domains are single lexical items. Artstein (2004) extends Rooth's alternative semantics to allow metrical feet to be marked as foci within phonological representations of words—the alternative set in this case contains two elements (Minnesota and Sarasota) that are of the form {X+/so.ta/}, where 'X' is some well-formed metrical foot.

The alternative semantics approach provides a good formal description of the constraints on contrastive focus marking, but it does not aim to explain *why* natural languages should have developed such a system. The current analysis provides one dimension along which it is possible to explain, using general principles of rational behavior and information transmission, why languages with prosodic flexibility like English, German and many others should converge on a system of prosodic accent placement which is analyzable with alternative semantics. We propose to model the selection of focused material as a signaling game (Lewis 1969) which can be iterated to select prominent domains within sentences, phrases and words. The solution to this game is a strategy for selecting and interpreting minimal chunks of "critical" material which should be focused in order to minimize the effect of noise on communication.

We emphasize that our chief desideratum is an explanatory model of a single facet of information structure—the placement of prosodic foci, often called con-

[1] Note that the focus on "American" is optional, but the focus on "Canadian" is not, a fact which we'll address in 3.2.

trastive foci. The terminology surrounding these matters is notoriously ambiguous. Therefore, before fleshing out our game-theoretic approach, it is important to clarify exactly what is to be modeled, and most crucially, what is not to be modeled. Namely, the remainder of this section argues for a separation between contrastive focus, which we simply call focus, and a distinct but related information-structural phenomenon of givenness marking (Féry & Samek-Lodovici 2006, Selkirk 2007, Stevens 2013: 2014). *Contra* some recent accounts (Büring 2008, Wagner 2012b), we argue that a conflation of these two phenomena only serves to confuse the issue and lend unnecessary mystery to observations about accent placement.

1.2 A working definition of focus

It is clear that the interface between prosody and meaning relies on more than just an alternative-based notion of focus. For example, (9), taken from Ladd (1996, p.175), illustrates the phenomenon of "de-accenting". Ladd and others have noticed that sometimes contextually salient parts of an utterance lack prosodic prominence, even if the QUD predicts some accent. As Swerts et al. (2002, p.630) puts it, a de-accented element "might otherwise be expected to be accented."

(9) She gave me a German book, but I don't READ <u>German</u>.

In this example the direct object *German* is *de-accented* (indicated by underlining) not because it is common to any set of alternative utterances or answers to a question under discussion, but rather because its denotation is salient in the prior discourse. This shifts the default right-edge prominence onto the verb, even though the verb need not contrast with another verb meaning. Following Schwarzschild (1999), the direct object in this case is *given.* Under Schwarzschild's definition the object is given in virtue of the fact that there is a salient proposition in the discourse context—'she gave me a German book'—which entails the existential closure of the noun 'German', i.e. $\exists x.German(x)$. Prince (1981) points out that the term *given* is at least three ways ambiguous in prior literature on information structure. Using her taxonomy we could refer to Schwarzschild's givenness as givenness$_S$, where the subscript S stands for 'salience', though we stick with the simpler term. If we follow Stalnaker (1979) in imagining discourse context as a common ground containing shared knowledge propositions, then we may say that the DP *German* in (9) is given because there is a salient subset of the common ground which entails the existence of the German language. Here, salience can be taken as the property of being in the consciousness of the interlocutors, to echo Chafe (1974).

There are surely other semantic/pragmatic factors for determining accent placement, but much of the data regarding the connection between prosody and meaning in languages like English can be accounted for with some combination of (contrastive) focus and givenness. We tentatively define these two concepts as follows.

(10) i. A constituent, word, morpheme or word part C is **FOCUSED** iff C is the minimal identifying material that differentiates the denotation of the entire uttered phrase or word containing C from some contextually available set of contrasting alternatives (possibly competing answers to a question under discussion)

ii. FOCUSED Cs UNDERGO PITCH ACCENT ASSIGNMENT

(11) i. A syntactic phrase XP is **GIVEN** iff it has a salient antecedent in the discourse, i.e. if a salient subset of the common ground entails the existential closure of XP's denotation

ii. GIVEN XPs CAN BE DE-ACCENTED, BLOCKING STRONG PITCH ACCENT, EVEN IF IN FOCUS

Note the somewhat vague formulation that given XPs *can* be de-accented even if in focus. It is not the case the given XPs must be de-accented. For example, in all-given sentences, something must nonetheless bear accent. See Stevens (2013: 2014) for further discussion of the interaction of givenness and accent placement.

The phenomena of focus and givenness are closely related in that there is considerable overlap in their effect on pitch accent placement. In fact, many cases are ambiguous between the two. For example, one might ask whether the subject in the answer in (4) receives strong stress because it selects one from a set of contrasting possible answers, or whether the predicate *went to the movies* is marked as given due to the salience of the previous question, leaving the subject as the only possible bearer of accent. For this reason, some work has attempted to provide a unified generalization which collapses the two notions (see Wagner 2012b). However, much work has relied on maintaining a distinction (e.g. Féry & Samek-Lodovici 2006, Selkirk 2007, Beaver & Clark 2008), and in general, the data are easier to analyze if we do so. Stevens (2013: 2014) highlight two defining features of focus which differentiate it from givenness: the requirement of a contrastive interpretation, and the extensibility of focus marking to sub-phrasal and sub-lexical domains. A definition of what counts as "contrastive" can be tentatively borrowed and adapted from Wagner (2012b):

(12) An element C *contrasts* with an element C' of the same semantic type iff the universal closure of C entails the falsity of the existential closure of C' (under common contextual assumptions).

That is, {'American farmer','Canadian farmer'} is a valid set of contrasting alternatives for the purposes of focus assignment iff the universal closure of 'American farmer', $\forall x.American(x)\ \&\ farmer(x)$ entails the falsity of the existential closure of 'Canadian farmer', $\exists x.Canadian(x)\ \&\ farmer(x)$, and vice versa. This holds under the common contextual assumption that people are typically identified by a single nationality. Assuming such a typical case, American farmers cannot also be Canadian farmers. Perhaps more intuitively, we may think of a contrastive relationship as an attribute-value relationship: two denotations contrast if they indicate attribute values along the same set of semantic dimensions. So any NP denotations selecting values for the attributes {NATIONALITY, PROFESSION} contrast with each other, e.g. 'American farmer' with 'Canadian carpenter', but not with 'short bachelor' or 'disagreeable extrovert'. In the case of QUDs, we may say that alternative answers contrast with each other when only one answer is the "correct" answer given the context. For example (4) above, if both interlocutors assume that the speaker has full knowledge of the topic at hand and wants to provide a complete answer, the set of contrasting possible QUD answers must be something like {'(only) Bob went to the movies','(only) Alice went to the movies','(both) Bob and Alice went to the movies',$\cdots$}.

Where this notion of contrast sets the pragmatic interpretation of focus apart from that of givenness, the two notions are differentiated structurally as well. Namely, the distribution of givenness marking (with no contrastive alternative set required) seems to be limited to syntactic XPs which are maximal projections, i.e. not immediately dominated by an XP of the same type. Borrowing loosely from Beaver & Clark (2008), we may say that XPs with highly discourse-salient meanings can become "inactive" for prosodic accent assignment, independently of whether they are focus-marked. But salience alone is not sufficient to license de-accenting within non-maximal sub-phrases or within individual words; in those cases, non-default accent patterns can only arise as a reflex of contrastively interpreted focus. The following examples illustrate.

(13) a. I didn't buy [MYSELF]$_F$ a convertible, I bought [my NIECE]$_F$ a convertible (contrastive focus attracts prominence)
b. If you hate convertibles so much, then why did you buy your NIECE <u>a convertible</u>? (DP de-accented due to givenness)

(14) a. It was a [RED]$_F$ convertible, not a [BLUE]$_F$ convertible (contrastive focus attracts prominence)
b. #If you hate convertibles so much, then why did you buy your niece a RED <u>convertible</u>? (given sub-NP cannot be de-accented)
c. If you hate convertibles so much, then why did you buy your niece a red CONVERTIBLE? (default prominence pattern in spite of givenness)

(15) a. It wasn't competitively priced, it was [OVER]$_F$priced (contrastive focus attracts prominence)
b. ??Speaking of price, how OVER<u>priced</u> was your CONVERTIBLE? (given morpheme cannot be de-accented)
c. Speaking of price, how overPRICED was your CONVERTIBLE? (default prominence pattern in spite of givenness)

Note that in (14-b), pitch accent on *red* would be possible if instead of the full noun *convertible* we used the anaphor *one*. But this can be seen as a general phonological fact about anaphoric pronouns, a reflex of the fact that there is default phonological pressure for closed-class function words to carry no accent (German et al. 2006), making the adjective a better candidate to bear the default prominence. The crucial fact is, when we use the full noun, which is felicitous when that noun is stressed, the de-accenting is blocked. These restrictions on the distribution of salience-motivated de-accenting may point to a particular syntactic analysis of givenness (Stevens 2014), but for current purposes it suffices to acknowledge that there is a real difference between the two phenomena. To assume otherwise would place an unnecessary burden on our model, e.g. by requiring an explanation of why (13-b) is acceptable but not (14-b). Now, having clearly delineated the scope of what we are analyzing, we begin introducing the current model by first summarizing the relevant game-theoretic notions.

2 Games, signals and equilibrium

This section provides a brief primer on the game-theoretic concepts necessary to understand signaling game models.

2.1 Basics

Game theory is a simple mathematical framework within which to model strategic interaction. It formally describes and predicts the behavior of two or more agents in interactive environments where each agent's decisions potentially affect the outcomes of the others' decisions. Depending on the players' strategies (i.e. the moves they make), the outcome of a game is better or worse for each player according to their preferences. Preferences about the outcome of a game are encoded in a payoff function called a utility function. Generally speaking, rational players should behave so as to maximize their utility as best they can.

A game can have any number of players, but for our purposes we only consider two-player games. For now, we can simply label the players 1 and 2. In the simplest case, each player has a single set of *actions* available to them—these can be thought of as the possible moves in the game. Let A_1 be the set of actions available to player 1 and let A_2 be the set of actions available to player 2. For a given sequence of actions, there is a particular payout or *utility* given to each player. In a simple game in which both players make only one move each, we can define a utility function U_n for player n as a function from $A_1 \times A_2$ to the set of real numbers, where player n prefers an outcome with a higher value for U_n to an outcome with a lower value. In other words, each player's goal is to maximize their own utility, and to do so they must consider the possible actions of the other players as well as their own actions. A single-move two-player game can thus be formalized as a tuple $\langle\{1,2\}, A_1, A_2, U_1, U_2\rangle$: two players, a set of actions available to each player, and a utility function for each player.

To maximize one's own utility requires not only a careful consideration of one's own actions, but also a consideration of how the other player will behave, since utility is determined by the sequence of actions taken by both players. Each player generates a *strategy*, which is a complete specification of which action is optimal in any possible state of gameplay. The outcome of a game played between perfectly strategic players is determined by assuming rational, utility-optimizing decision making from each player at each step of play. The result of such reasoning is an outcome where neither player has any reason to unilaterally deviate from her strategy. Such an outcome is called an *equilibrium*, and an individual player's strategy which leads to equilibrium is an *equilibrium strategy*.

Much of the early work in game theory was devoted to games in which one player can gain utility at the expense of the other. For example, consider the variant of a "split or steal" game (SoS) represented in Fig.2. The two rows correspond to the actions available to player 1, who we can call the "row player", or R; the columns correspond to those of player 2, the "column player", or C. As per the above-given formalization of such games, we can represent SoS as a tuple $\langle\{R,C\}, A_R = \{\text{Split}, \text{Steal}\}, A_C = \{\text{Split}, \text{Steal}\}, U_R, U_C\rangle$, where the utility functions U_R and U_C are specified by the first and second numerals, respectively, in each cell of Fig.2. In plain English, "split or steal" is a game where two players are given an amount of money—\$100 in this case if we assume that the utility functions correspond to raw monetary payouts—and both players are given an option to choose to "split" or to "steal". If both players choose to split, the cash is split 50/50 between the two players. But if one player chooses to steal, that player gets the entire amount, leaving the other player with nothing. The catch

	Split	Steal
Split	50,50	0,100
Steal	100,0	1,1

Fig. 2 A simple two-player game

is, if both players choose to steal, then the house keeps the prize money, and the players walk away with a small consolation prize of $1.

The players would collectively do better with a "Split/Split" outcome than with a "Steal/Steal" outcome; however, both players individually have incentive to play "Steal". Equilibrium is calculated such that each player considers only her own best move for each possible action taken by the other player. Both players will reason the same way: if my opponent plays "Split", my utility is maximized by playing "Steal", and if my opponent plays "Steal", my utility is maximized by playing "Steal". Both players reason that the best strategy is to steal, and the outcome of the game is the consolation prize outcome. Formally, this outcome meets the requirements for a *pure strategy Nash equilibrium*, after Nash (1950: 1951). The requirements can be specified as follows, where the equilibrium state is labeled $\langle a_R^*, a_C^* \rangle$.

$$\begin{aligned} \forall a \in A_R : U_R(\langle a_R^*, a_C^* \rangle) \geq U_R(\langle a, a_C^* \rangle) \\ \forall a \in A_C : U_C(\langle a_R^*, a_C^* \rangle) \geq U_C(\langle a_R^*, a \rangle) \end{aligned} \tag{1}$$

In games like SoS, selfish reasoning leads to outcomes that are less optimal for the players than outcomes obtained by reasoning selflessly. The equilibrium concept outlined above is limited in that it alone does not account for behavior in games where players have common goals, or in games where players can communicate. So we now turn to some enrichments and refinements that have been made over the years to account for these phenomena, which ultimately allow us to specify a particular class of signaling games which utilize these concepts.

2.2 Coordination

Imagine a scenario where two friends, Robert and Charlotte, are leaving their houses to go meet each other at a pre-determined time and place for a business meeting. Both friends had previously agreed to meet at noon at "the best cheesesteak restaurant in Philadelphia." Robert realizes en route that, while he believes Pat's to be the best cheesesteak restaurant in Philadelphia, Charlotte once mentioned that she believes Geno's to be the best cheesesteak restaurant in Philadelphia. At the same time, Charlotte comes to a similar realization about Robert's cheesesteak preference. A dilemma arises: at which location should the friends meet? This dilemma can be represented as a simple game where the utility for each of the four possible outcomes corresponds to a binary payout: 1 for the preferred outcome of meeting at the same place and 0 for the dispreferred outcome of going to different locations. We can represent this as $\langle \{R, C\}, A = \{\text{Pat's}, \text{Geno's}\}, U_R, U_C \rangle$, where R(obert) is the row player and C(harlotte) the column player, and where the utility functions U_R and U_C, which happen to be identical in this case, are represented in Fig.3.

	Pat's	Geno's
Pat's	1,1	0,0
Geno's	0,0	1,1

Fig. 3 A simple coordination game

This game instantiates a class of games first discussed by Schelling (1960); it is a *coordination game*. In such a game the players must match their behaviors by choosing the same action from a common set of possible actions. The utility functions in our coordination game are such that if the players do not take the same action, utility for both players is zero.

The conditions for Nash equilibrium given above (with A_R and A_C set equal to A) hold for both the coordinating outcomes—there are two such equilibria. The problem that R and C are confronted with is the problem of selecting the same one. In the absence of any further guidance, the best the players can do is choose randomly, and fail to coordinate half of the time. But in reality, Robert and Charlotte are likely to have *beliefs* about the likelihood of one outcome or another. For example, imagine that Robert was once forcibly removed from Geno's for insulting an employee, whose name is Fred. It would be reasonable for Charlotte to believe that Robert does not want to go back there if that same employee is working, for fear of further embarrassment. We can represent this belief as a probability function. Say that C believes that with some positive probability P_{Fred}, Fred is currently at work, in which case Robert will definitely end up at Pat's. If Fred is not working, then the probability of Robert choosing one action or the other is believed to be $\frac{1}{2}$. We can calculate C's *expected utility* for choosing Pat's vs. Geno's, where C's expected utility is the probabilistically weighted average of her utility for all possible outcomes given a particular choice of action.

$$\begin{aligned} EU_C(\text{Pat's}) &= P_{\text{Fred}} \cdot 1 + (1 - P_{\text{Fred}}) \cdot \frac{1}{2} \\ EU_C(\text{Geno's}) &= P_{\text{Fred}} \cdot 0 + (1 - P_{\text{Fred}}) \cdot \frac{1}{2} \end{aligned} \tag{2}$$

If Charlotte wants to maximize her expected utility given her beliefs—which is the best she can do given her uncertainty about Robert's action—then she does best to choose Pat's. If it is common knowledge that C has these beliefs, then R knows that C, if she is a rational reasoner, will choose Pat's. In that case, R should assign a probability of 0 to C choosing Geno's, giving that action zero expected utility for R. So R should choose Pat's, too. Finally, if C knows that R will choose Pat's, then she knows that her strategy will pay off, and she has no reason to deviate. So we end with a pair of optimal strategies where both players choose to go to Pat's, regardless of their initial opinion about cheesesteak quality, and therefore successfully coordinate. This pair of strategies was obtained by maximizing expected utility for each player given their beliefs about the other player's behavior. An equilibrium obtained by reasoning about expected utility based on rational and consistent beliefs about the state of the game is called a *perfect Bayesian equilibrium* (PBE) (Harsanyi 1968, Fudenberg & Tirole 1991). This is the equilibrium concept relevant to the class of signaling games, where PBEs can be generated for signaling games using Iterated Best Response (IBR) reasoning (Franke 2011). The remainder of this section outlines these concepts, which form the basis of our model of focus selection.

2.3 Signaling

The Pat's/Geno's coordination game becomes trivial if one player can simply send a text message to the other player and inform them of their intended behavior. Robert decides which restaurant to go to, sends a text message to Charlotte, either "I'll be at Pat's" or "I'll be at Geno's", and after receiving the message, Charlotte decides where to go, and both players receive positive utility iff they choose the same place. This can be modeled as a *signaling game* (Lewis 1969). In these games, there are two players, and Sender and a Receiver. The Sender is probabilistically assigned a *type* prior to play (Pat's vs. Geno's), the Sender then sends a *message* to the Receiver (the text message), which has some conventional semantic content, and the Receiver then takes an *action* (Pat's vs. Geno's). Utility functions for both players can depend on the Sender's type, the Receiver's action, and the message that was transmitted. Simply put, Robert's type is the information he wants to convey to Charlotte, and Charlotte's action is the information that she thinks that Robert wants to convey.

More formally, we represent a signaling game as a tuple of the following form:

$$\langle \{S, R\}, \Phi, M, [\![\cdot]\!], T, \delta, A, U_S, U_R, \mathcal{C} \rangle$$

where S and R are the Sender and Receiver, respectively, Φ is a set of well-formed semantic formulae, M is a *language* of possible messages, $[\![\cdot]\!]$ is a denotation function from M to Φ, T is the set of possible Sender types, δ is a prior probability distribution over those types, A is the set of possible Receiver actions, U_S and U_R are the utility functions for Sender and Receiver, respectively, the first a function from $T \times A$ to the natural numbers, the second from $T \times M \times A$ to the natural numbers, and finally, $\mathcal{C}$ is a *cost* function from $M \times T$ to the natural numbers, encoding possible reductions in Sender utility for more effortful messages.

We consider signaling games where T is identical to A, where both contain members of Φ as well as a special type/action known as the *babbling* type/action, which we will indicate with the symbol '#', such that T and A are both subsets of $\Phi \cup \{\#\}$. The babbling Sender, $t_\#$, is a low-probability Sender type that represents the possibility of a Sender who does not obey Grice's Cooperative Principle. As such, $t_\#$ can in principle send any message without regard for meaning or efficiency. The inclusion of $t_\#$ and its corresponding response $a_\#$ makes for a cleaner and more predictive model. Proper equilibria in Bayesian games such as these require that the Receiver have a rational response to "off-equilibrium" messages. This means that if the Receiver receives a message that could not have been optimal for any cooperative Sender type, we still need to specify how the Receiver should act; otherwise our model is incomplete. We hold that the Receiver in such cases should assume, in the absence of some repair that can be made to the game to explain the Sender's behavior (see Section 5), that the Sender is behaving irrationally. The inclusion of $t_\#$ encodes this, where $t_\#$ has the following properties: (i) $t_\#$ "babbles" in the sense that $t_\#$ can in principle send any message, i.e., cares neither about meaning nor about efficiency, and (ii) $t_\#$ nonetheless wants to signal the correct type, i.e., is not trying to trick the Receiver into believing that any meaning is intended. Given the Cooperative Principle, the Receiver should assume *a priori* that the Sender is non-babbling, i.e. that the prior probability of $t_\#$, $\delta(t_\#)$, is very close to zero. This assumption can be built into the IBR mechanism by having the Receiver only consider $t_\#$ for messages that could not have been sent by a non-babbling type. (We might call this "rationality *ceteris paribus*.") The result of this is that any message that can't be an equilibrium best response for a non-babbling type will be taken as an equilibrium best response for $t_{\#}$, and the Receiver will take the corresponding action $a_{\#}$. For the simpler cases, this completely erases the problem of responding to off-equilibrium messages (because there aren't any), and gives a very simple way to define what counts as felicitous or infelicitous: When we apply our signaling model to focus placement, foci that correspond to messages which can only be sent by $t_{\#}$ result in infelicitous prosodic patterns.

To put it slightly differently, in these signaling games, the Receiver assumes *a priori* that the Sender wants to convey some meaning with his message, only considering the possibility of a babbling Sender when no other types are possible. A non-babbling Sender's type is conceived of as the meaning that that Sender wishes to convey (some semantic formula in Φ), and the Receiver's action is conceived of as the Receiver's best guess as to S's type, i.e. the meaning which the Receiver assumes is being conveyed. The signaling games which we are concerned with are also games of coordination insofar as we assume that positive utility is rewarded to both players only if S's type is equal to R's action. For the current example, let P denote the proposition 'Sender will be at Pat's.' If S is of type t_P, he wants to convey the proposition 'S will be at Pat's', and the action a_P corresponds to R attributing that same meaning to S (and therefore, we assume, deciding to go to Pat's). Both the type and action labels serve simply as stand-ins for the semantic formulae which they pick out, which are represented by the subscripts in those labels. Under this notational convention, coordination occurs iff $i = j$ for outcome $\langle t_i, m, a_j \rangle$, where m is whichever message from M that S decided to send to R. This includes the outcome where $i = j = \#$.

Encoding a preference for shorter messages for non-babbling types, we assume that, for those types, the cost function $\mathcal{C}(m,t)$ increases with the length of message m, crudely construed as the number of syllables in m, and that this is subtracted from S's base payoff for coordinating. $\mathcal{C}(m,t)$ is assumed to be 0 for babblers, who are not concerned with message efficiency. The utility functions are formulated as follows.

$$\begin{aligned} U_R(t,a) &= 1 \text{ iff } a = t \\ &= 0 \text{ otherwise} \end{aligned} \tag{3}$$

$$U_S(t,m,a) = U_R(t,a) - \mathcal{C}(m,t) \tag{4}$$

The Receiver's utility depends only on whether coordination occurs, i.e. it is independent of the content of S's message. However, we use a variant of a signaling game which Franke (2011) calls an *interpretation game* in which R has a *default belief* that S's message conveys truthful semantic content. That is, R begins with the belief that $[\![m]\!]$ is a truthful proposition, and only alters this belief if R reasons that S is either a babbler, or else has some incentive to be either dishonest or non-literal. An interpretation game representation of the Pat's/Geno's game is depicted in Fig.4, with only raw utilities given (i.e. without considering the cost term for the Sender). Assuming the costs to be very small, we can simply incorporate them into the reasoning procedure used to derive an equilibrium—we simply assume that, all things being equal, S always chooses a shorter message over a longer one. There are two meaningful types in Fig.4: either S wants to signal Pat's (t_P) or else Geno's (t_G); accordingly, there are two possible actions for R to take, a_P and a_G. S may send a text message with semantic content indicating that he

	a_P	a_G	$a_\#$	$m_\emptyset$	m_P	m_G
t_P	1,1	0,0	0,0	✓*	✓*	✗
t_G	0,0	1,1	0,0	✓*	✗	✓*
$t_\#$	0,0	0,0	1,1	✓	✓	✓

Fig. 4 A simple coordination game with meaningful signaling

will be at either Pat's (m_P) or Geno's (m_G). We also consider the possibility of a null message, i.e. a message of length zero.

Each possible message is given a column, under which a check mark in the row corresponding to a particular type denotes that that message is compatible with that type given R's default beliefs, and an 'x' denotes incompatibility. In this context, a message m is 'compatible' with a type t given a belief b iff the probability of a type-t Sender sending message m is non-zero assuming b. The null message is compatible with any type, and all messages are compatible with $t_\#$. Finally, an asterisk next to a check mark indicates that the type associated with that row is maximally probable compared to other types which are compatible with the message associated with that column.

Franke's proposed method of equilibrium generation for interpretation games, the *iterated best response* (IBR) method, is a formal procedure for reasoning strategically about expected utility, and in the case that the IBR procedure converges on a single optimal strategy for each player, that set of strategies is a perfect Bayesian equilibrium. Simply put, the players reason about what Sender types do best by sending which messages, such that they generate a set of optimal type-message pairs. If this set of type-message pairs is sufficient for the Receiver to use the Sender's message to guess the Sender's type with complete confidence, then equilibrium is guaranteed.

This reasoning proceeds hierarchically à la the "level-n reasoning" approach of Camerer et al. (2004) and Bardsley et al. (2010), but it is idealized in that the number of levels is in theory unbounded. We begin with a "level-0" Receiver (R_0)—a Receiver that takes all messages at face value, i.e. assumes that the Sender is always non-babbling, honest and literal. We calculate, for each possible message, which Sender type(s) an R_0 will assume (which determines what action the Receiver will take). We then calculate, for each possible type, which message a "level-1" Sender (S_1)—a Sender who assumes that the Receiver is an R_0—would send in order to maximize his own expected utility. (That is, S_1 maximizes the probability of successful coordination given how R_0 will respond to each message.) Then, we consider, for each possible message, which Sender type(s) a "level-2" Receiver (R_2), who assumes the Sender is an S_1, would assume given that message. These levels of reasoning continue until the Speaker's and the Receiver's reasoning converge on an equilibrium. If there is a unique optimal message for every possible non-babbling type, and if the set of possible messages produced by the babbler is non-overlapping with the set of potentially optimal messages for non-babblers (meaning the Receiver can always correctly assess the Sender's type from their message), then the equilibrium is a *separating equilibrium.*

More formally, a level-n Sender of type t selects a message m which maximizes the conditional probability $\text{Prob}(t|R_{n-1}(m))$, defined as follows, where $R_{n-1}(m)$ is the (possibly singleton) set of Sender types that a level $n-1$ Receiver assumes to be compatible with m.

$$\begin{aligned}\text{Prob}(t|R_{n-1}(m)) &= \frac{1}{|R_{n-1}(m)|} \text{ if } R_{n-1}(m) \text{ contains } t \\ &= 0 \text{ otherwise}\end{aligned} \tag{5}$$

In other words, S is trying to correctly signal his type, and thus wants to maximize the probability that his message will lead R to the belief that S is indeed of that type. A level $n+1$ Receiver, then, pairs each possible message with the type or set of types that could have been sent by a level n Sender, and so on.

For the game in Fig.4, we obtain an equilibrium in the following way. First, we adopt the following notational conventions. $R_n \Rightarrow \{m_A, m_B \rightarrow \{t_1, t_2\}\}$ means, 'a level-n Receiver interprets both message m_A and message m_B as being compatible with either type t_1 or t_2', while $S_n \Rightarrow \{t_1, t_2 \rightarrow \{m_A, m_B\}\}$ means, 'a level-n Sender of either type t_1 or type t_2 maximizes the likelihood of successful coordination by sending either message m_A or message m_B.' We then need to go through four steps of the procedure (levels 0 to 3) before convergence occurs.

$$\begin{aligned}
R_0 &\Rightarrow \{m_P \rightarrow \{t_P\}, m_G \rightarrow \{t_G\}, m_\emptyset \rightarrow \{t_P, t_G\}\} \\
S_1 &\Rightarrow \{t_P \rightarrow \{m_P\}, t_G \rightarrow \{m_G\}, t_\# \rightarrow \{m_P, m_G, m_\emptyset\}\} \\
R_2 &\Rightarrow \{m_P \rightarrow \{t_P\}, m_G \rightarrow \{t_G\}, m_\emptyset \rightarrow \{t_\#\}\} \\
S_3 &\Rightarrow \{t_P \rightarrow \{m_P\}, t_G \rightarrow \{m_G\}, t_\# \rightarrow \{m_\emptyset\}\} \\
\text{Eq.} &\Rightarrow \{\langle t_P, m_P, a_P\rangle, \langle t_G, m_G, a_G\rangle, \langle\ t_\#, m_\emptyset, a_\#\rangle\}
\end{aligned} \tag{6}$$

Since R_0 can interpret the null message as being sent by either m_P or m_G, S_1 excludes that message for non-babbling types, even though it is less costly, on the grounds that it does not maximize $\text{Prob}(\cdot|R_0(m))$.[2] Knowing this, R_2 will interpret a null message as a signal of $t_\#$, the only S_1 type who could have sent that message. After one more step, S and R have converged on a pairing between types $t_P/t_G/t_\#$ and messages $m_P/m_G/m_\emptyset$, respectively. For each message, there is a unique action on the part of R which will guarantee successful coordination, and thus a separating equilibrium has been reached.

2.4 Incorporating partial signals and salience

So far we have assumed that non-null messages correspond to propositions, and that the Receiver's default belief is that the proposition conveyed by a non-null m is true. However, to model the selection of prosodic foci within a sentence we will also need to consider the possibility of "partial" signals, e.g. messages whose denotations are sub-propositional, but which are being used to signal a propositional meaning. For example, in the game above, we might add two additional possible messages: instead of sending a full-sentence text message "I'll be at Pat's/Geno's",

[2] This carries the assumption that R's default response to type uncertainty is random selection. One could imagine a strategy where R always responds to a null message with a_P, for example, even if both types are equiprobable. But in the absence of some principled reason to prefer one action over another, the only natural assumption is there is a 50/50 chance of either action when R is wholly uncertain about S's type. Moreover, R does not benefit in any way from deviating from a random response: the probability of success is 0.5 in any case.

S may send a simple one-word message, either "Pat's" or "Geno's". Logically, the entire proposition conveyed by "Pat's" could be something like 'Pat's was opened in 1930'. However, the context in which the game is played dictates that the Sender is most likely indicating where he wants to meet, since that is a the problem at hand. We must encode this in the game by restricting the set of types/actions such that S always wants to signal (and R always wants to interpret) a relevant proposition of the form 'I'll be at X'. This restriction is taken to be a natural consequence of the context in which the message is sent.

In this slightly revised version of the interpretation game, which we will call a "focus game" in the context of this paper, we must adjust our formulation of R's default belief. In a focus game, R starts with the belief that $[\![m]\!]$ is a *component meaning* of S's type. In other words, R begins with the assumption that the Sender is signaling information about his type in the form of a partial message. Intuitively, the assumption is that the meaning of the transmitted message is "contained within" the larger intended meaning. Informally, we may assume the following definition.

(16) For two semantic formulae A and B, A is a component meaning of B iff A composes with some semantic formula C to produce B, where C does not contain any vacuous binding (e.g. $\lambda x.teach(sue, phonetics)$).

Formally, we can recursively define a relation $\rightarrow_{comp}$ with respect to a language M. By anchoring the definition of component meaning-hood to M, we can exclude vacuous formulae like $\lambda x.teach(sue, phonetics)$ in virtue of such formulae not being conveyable in any reasonable characterization of M. In general, restricting M also restricts the set of component meanings under consideration, which has positive consequences for the model, as we will see in 3.2.

$$A \rightarrow_{comp} B \text{ in } M \text{ iff either: } \begin{array}{l} \text{(i) } \exists \rho \in M : [A([\![\rho]\!]) = B] \vee [[\![\rho]\!](A) = B], \text{ or} \\ \text{(ii) } \exists C.A \rightarrow_{comp} C \ \& \ C \rightarrow_{comp} B, \text{ or} \\ \text{(iii) } A = B \end{array} \tag{7}$$

For the base case, A is a component meaning of B with respect to M if there is an expression in M whose denotation can combine with A to produce B. Moreover, if A is a component meaning of a component meaning of B, then A is a component meaning of B (i.e. component meaning-hood is transitive). Finally, component meaninghood is reflexive. With this we can factor a proposition ϕ into its component semantic parts. The proposition $teach(sue, phonetics)$ is factored as follows.

$teach(sue, phonetics)$
sue
$\lambda x.teach(x, phonetics)$
$phonetics$
$\lambda x.\lambda y.teach(x, y)$
$teach(sue, phonetics)$

Messages with one of these denotations are deemed by R_0 to be "compatible" with a (non-babbling) Sender whose type is the intended proposition, 'Sue is teaching phonetics.' This default assumption allows, for example, the Receiver to correctly interpret the message "Pat's" as signaling the proposition 'I'll be at Pat's', because

	a_P	a_G	$a_\#$	$m_\emptyset$	m_p	m_g	m_P	m_G
t_P	1,1	0,0	0,0	✓*	✓*	✗	✓*	✗
t_G	0,0	1,1	0,0	✓	✗	✓*	✗	✓*
$t_\#$	0,0	0,0	1,1	✓	✓	✓	✓	✓

Fig. 5 Interpretation game with partial signals and non-flat priors

the context of the game dictates that that proposition is the only possible type for which the denotation of "Pat's" is a component meaning.

Finally, we have so far been assuming equiprobability of non-babbling types (given by the δ function in our game specification). That is, the above examples assume that there is no *a priori* reason to believe that one intended meaning is more likely than another. In reality, however, as pointed out in 2.2, the players' expected utility may be influenced by prior beliefs or biases. For current purposes, we posit that *salience* (discussed in more detail in Section 3) can bias the Receiver toward one type or another, such that otherwise ambiguous messages may tip the Receiver in a particular direction. If the Sender knows this, then he can exploit it to send less effortful messages. This is encoded in the game by raising the value of δ for more salient types.

The role of δ can be incorporated into the IBR procedure in a way similar to cost: we simply assume that, in the case of multiple compatible types given a particular message, the Receiver only considers those which maximize the prior probability, $\delta(t)$. Note that, for all of the examples we are interested in, this entails our assumption above that R_0 does not consider $t_\#$, because $t_\#$ is taken to be inherently low-probability, and there is always a meaningful type available for any message at R_0. Maximizing δ at each step ensures that, whichever type(s) R assumes given S's message, R's assumption is motivated by the maximization of expected utility, as defined below.

$$
\begin{gathered}
EU_R(a|m) = \sum_{t \in T} P(t|m) U_R(t,a) \\
\text{where, by Bayes' theorem, } P(t|m) \propto \delta(t) \cdot P(m|t)
\end{gathered}
\tag{8}
$$

Putting it all together, we consider an extension of our Pat's/Geno's signaling game in Fig.5. Let $m_{P/G}$ be full-sentence messages, "I'll be at Pat's/Geno's", and let $m_{p/g}$ be partial messages, "Pat's" or "Geno's". Again, we consider a null message. Messages are arranged in increasing order of effort, as indicated by double lines. Maximal δ values for compatible types are indicated with asterisks as before, but this time, δ is taken to incorporate salience-based biases. Fig.5 assumes that (it is common knowledge that) the Receiver has a prior bias toward assuming that S is of type t_P.

To find equilibrium for this game, we incorporate partial signals and prior probabilities into our final variant of the IBR procedure, outlined below.

1. R_0:
   For all m in M, output the most probable type(s) "compatible" with m:
   - Partially order T from highest to lowest probability according to δ.
   - Output $T^*(m)$, where $T^*(m)$ is the set containing the highest-ranked type(s) t such that $[\![m]\!] \rightarrow_{comp} t$.
2. S_1:
   For all t in T, output the best message(s) to send to R_0:

– Calculate $M^*(t) \subseteq M$ such that for all m in $M^*(t)$, $\text{Prob}(t|R_0(m)) \geq \text{Prob}(t|R_0(m'))$ for all $m' \in M$.
– Partially order $M^*(t)$ by effort (from fewest to greatest number of syllables).
– Output the set containing the lowest-effort message(s) in $M^*(t)$.

3. R_2:
For all m in M, output the most probable type(s) that do best to send m to S_1:
– Output the set containing the δ-maximizing type(s) t such that $m \in S_1(t)$.
4. For $n \in \{4, 6, 8, \cdots\}$, calculate S_{n-1} and R_n by analogy to S_{n-3} and R_{n-2}, respectively, until convergence occurs.
5. Convergence occurs at a level S_n when for any given type/message pair $\langle t, m \rangle \in T \times M$, $t \in R_{n-1}(m) \leftrightarrow m \in S_n(t)$.
6. If each tuple $\langle t, m, a \rangle$, where $a \in R_{n-1}(m)$ and $m \in S_n(t)$, maps a single type to a distinct (set of) value(s) for m, then the set containing those tuples is a separating equilibrium.

We now apply this procedure to Fig.5.

$$\begin{aligned}
R_0 &\Rightarrow \{m_\emptyset, m_p, m_P \to \{t_P\}, m_g, m_G \to \{t_G\}\} \\
S_1 &\Rightarrow \{t_P \to \{m_\emptyset\}, t_G \to \{m_g\}, t_\# \to M\} \\
R_2 &\Rightarrow \{m_\emptyset \to \{t_P\}, m_g \to \{t_G\}, m_p, m_P, m_G \to \{t_\#\}\} \\
S_3 &\Rightarrow \{t_P \to \{m_\emptyset\}, t_G \to \{m_g\}, t_\# \to \{m_p, m_P, m_G\}\} \\
\text{Eq.} &\Rightarrow \{\langle t_P, m_\emptyset, a_P \rangle, \langle t_G, m_g, a_G \rangle, \langle\ t_\#, \{m_p, m_P, m_G\}, a_\# \rangle\}
\end{aligned} \tag{9}$$

We obtain a separating equilibrium. If S is of type t_P, he can exploit the prior bias toward his type, and say nothing at all. S's silence is interpreted as a signal that the default biases hold: $m_\emptyset$ prompts a_P, and coordination occurs at minimal cost to the Sender. If S is of type t_G, the message m_g ("Geno's") is optimal, which prompts a_G, also guaranteeing coordination.

A note on our incorporation of $t_\#$ into this example, whose best response is any message from $\{m_p, m_P, m_G\}$: The '#' notation should not be taken too literally for this illustrative example. In this simple, one-off Pat's/Geno's game, it would not be seen as fully infelicitous to send the longer-than-needed messages associated here with $t_\#$. Rather, these messages are marked off in this game as being sub-optimal. In the next section, when we apply games of this type to focus marking, we will introduce the linking hypothesis that messages associated with $t_\#$ are not only sub-optimal, but also result in infelicitous prosodic patterns. It is by way of looking ahead to that case that we have chosen the '#' notation.

Now that we have introduced the relevant game mechanics, we propose a model using interpretation games with partial signals and salience-based priors to select informationally crucial material for purposes of prosodic focus assignment.

3 Focus games

In Section 1 we discussed the notion that the placement of prosodic focus can be conceived of as the selection of material to undergo prosodic accent assignment such that that material conveys information that is most crucial for the interpretation of the speaker's intended meaning. After Schmitz (2008), Stevens (2013), Bergen & Goodman (2014) and others, we may posit that such a system is, in a sense, a reasonable and rational system for languages with prosodic grammars that allow it to arise. This is because, adopting a "noisy channel" model of communication (Shannon 1948, Benz 2012), material that is in focus is less likely to be obscured by noise due to its increased perceptual prominence. Of course, there are intervening factors which prevent the picture from being so simple. Phonology dictates that the entire focus cannot be uniformly prominent—there must be peaks and valleys within a phrase of any substantial length—and the ability to prosodically "de-activate" contextually given XPs (Beaver & Clark 2008) can obfuscate the issue of whether prosodic prominences rely on the presence of contrasting alternatives à la Rooth (1992), Wagner (2012b) and others.

In 1.2 we motivated a definition whereby focus explicitly refers to the phenomenon of assigning prominence within domains which serve to identify one element from a set of contrasting alternatives. It is this phenomenon which is to be modeled as a process of selecting the minimal signal conveying crucial information, such that a perfectly rational hearer can reconstruct the speaker's intended meaning from that signal alone. In other words, we hold that focused elements alone allow the hearer to reconstruct larger chunks of meaning within constrained contexts by exploiting the assumption that meanings are to be selected from a well-defined and accessible set of possibilities.

This model is formalized as a signaling game of the kind outlined in 2.4, with the speaker as the Sender and the hearer as the Receiver, but with a few additional assumptions, all of which are independently motivated by general principles of behavioral game theory (e.g., by theories of team reasoning, see Bacharach 2006). Namely, all possible intended meanings (Sender types) and possible interpreted meanings (Receiver actions) are taken from a common set of *contextually available* alternatives, where contextual availability is defined below in terms of entailment under existential closure by CG_S (Schwarzschild 1999), where CG_S is the subset of the common ground consisting of the propositions that are salient in the sense of being assumed to be in the consciousness of the interlocutors (Chafe 1974):

(17) A set of alternatives Φ is contextually available iff:
  a. It is a salient fact that one of the members of Φ has a true existential closure. (informal)
  b. $CG_S \Rightarrow [\exists \phi \in \Phi.\ \mathrm{ExClo}(\phi)]$ (formal)

This operationalizes the idea that a set of alternatives must be "activated" by being evoked by the discourse context, either by being introduced into the context directly, as with a QUD, or by one of the set's members being introduced. In the former case, the condition in (17) holds because CG_S directly entails that there exists a true member of the QUD—its answer (which, being a proposition, is its own ExClo). Of course, the answer to the QUD need not be in the salient common ground, only that there is an answer. In the latter case, e.g. when a discourse antecedent like 'American farmer' evokes a 2x2 semantic space of nationality-profession pairs, CG_S entails the ExClo of the antecedent ('American farmers exist'), and thus trivially entails that there exists a true ExClo of a member of Φ. In sum, a set of alternative meanings is contextually available when it is saliently assumed that one of its members has an existential closure that is true.

Crucially, the members of Φ must lie along the same semantic dimension(s) (what Bacharach 2006 calls "a family of attributes"). This means that, for example, the semantic description 'blue convertible' is only possible as a type/action in the focus game if concepts of color and category are evoked via some other meaning, e.g. 'red sedan', that is salient in the current discourse context. If these semantic dimensions are not evoked by the context, then 'blue convertible' is not an available concept to be represented by the players as a possible move in the game. Moreover, if 'red sedan' is in fact salient, then every alternative color/category combination must be represented in the game. This cashes out the intuition of Bacharach & Bernasconi (1997, p.6) who, analyzing how players internally represent available actions in coordination games, remark that "when attributes do come to mind they come in clusters. . . it is nearly impossible to notice that 'U' is a vowel without noticing that other objects are consonants."

Prior probabilities for types in the focus game are determined by a scalar notion of discourse salience. The idea behind this is simply that some meanings can be more prominent within the consciousness of the interlocutors than others. For example, a prior mention of red convertibles will give some salience to the meaning 'blue convertible', but even more to 'red convertible', since both component meanings have been mentioned explicitly. This allows the hearer to have biases toward certain types, as we represented in 2.4. Finally, also as outlined in 2.4, the Receiver begins her equilibrium reasoning with the default belief that the Sender's message denotes a component meaning of his intended meaning (that is, a component meaning of his type).

Putting it all together, we are proposing an interpretation game to which the following constraints apply.

1. Non-babbling types and actions represent intended and interpreted meanings, respectively, both selected from a shared pool of formulae of the same semantic type.
2. The set of possible meanings is contextually available.
3. The space of possible meaningful Sender types creates a *partition* over an appropriate set of semantic objects. In other words, types are drawn from a set of mutually exclusive semantic descriptions, such that no two types could simultaneously serve as appropriate semantic contributions given the context.
4. The prior probability of a meaningful type is proportional to the salience of that type, construed as the degree of semantic overlap between that type and a discourse-salient default type .
5. The Receiver reasons from a default belief that the denotation of the Sender's message is a component meaning of the Sender's type.

Formally, where Φ_u is the set of uncurried set representations of the semantic formulae in Φ (e.g. a set of phonetics-teaching individuals for the predicate $\lambda x.teach(x, phonetics)$):

1. $A = T \subset \Phi \cup \{\#\}$
2. $CG_S \Rightarrow [\exists \phi \in \Phi.\ \mathrm{ExClo}(\phi)]$
3. $\forall \langle \phi, \phi' \rangle \in \Phi_u \times \Phi_u :\ \phi \neq \phi' \rightarrow \phi \cap \phi' = \{\}$
4. $\forall t \in T : t \in \Phi \rightarrow [\delta(t)\ \propto \mathrm{salience}(t)]$
5. R_0 assumes $P(t|m) > 0$ iff $[\![m]\!] \rightarrow_{comp} t$

Note that last of these constraints is not a constraint on game structure, but rather a constraint on the solution method that is applied to the game.

Also note that our third constraint—taken to be reflective of a general fact about how players represent a space of distinct options in a coordination game such as this (evoking the assumptions of Bacharach 2006, though in a slightly different context)—derives the definition of contrast from Wagner (2012b) which we tentatively adopted in 1.2. Wagner posits that contrasting alternatives have the property that the universal closure of one alternative entails the falsity of the existential closure of another alternative. By assuming a partition over the relevant meaning space, we constrain Φ, and by extension T/A, such that the characteristic set of any meaning in Φ must be non-overlapping with that of any other meaning in Φ. In a game selecting from among possible NP denotations, for example, this means that 'American farmer' is a valid alternative to 'Canadian carpenter' only if (given some contextual assumptions and restrictions), the set of American farmers and the set of Canadian carpenters is non-overlapping. If this is true, then the universal closure of one meaning, under which all entities are in the characteristic set for that meaning, is incompatible with the existential closure of the other, since no entity can be in both sets.[3]

As before, we include the possibility of a hypothetical *babbling Sender*, $t_{\#}$. This provides us with an easy way to determine infelicity with respect to focus marking in this framework: in the simplest case, if an utterance U contains a focus F with meaning $[\![F]\!]$ and the Receiver's optimal response to F in the appropriate focus game is anything other than $a_{[\![F]\!]}$, then U is infelicitous. This will be revised somewhat in Section 5 to account for higher-level pragmatic inferences that can result from "deviant" focus marking, but it is a useful idealization of the facts.

Having outlined the nature of focus games, we now propose the principal hypothesis underlying this model.

(18) Given an uttered sentence, phrase or word part C:
  a. A substring F of C is a focus if $\langle t, F, a\rangle$ is part of a separating equilibrium in a focus game in which t and a are equivalent to $[\![C]\!]$.
  b. If no such focus game is constructible given the discourse context, assign no foci within C.
  c. If an entire self-contained utterance U lacks foci, assign default prosody (de-accenting any given-marked constituents).
  d. Assign a strong pitch accent within each focus.

This hypothesis is a game-theoretic packaging of two simple principles: first, that focus is assigned to the most important or informative elements in a sentence, and second, that givenness and default prosodic rules also play a role in accent

[3] The partition requirement means that in the default case, QUD answers are taken to be exhaustive, e.g., "BILL is teaching phonetics" requires that Bill is the only relevant individual who teaches the relevant phonetics course, since the QUD answers must partition the space of possible worlds into non-overlapping propositions, which makes "BILL is teaching phonetics" incompatible with worlds in which both Bill and Sue are teaching phonetics. Cases involving partial knowledge or mention-some answers, e.g., where the speaker does not know or does not care whether Sue also teaches phonetics, can be accounted for by representing the QUD answers not as simple propositions about the world, but rather as propositions about epistemic states. For example "BILL is teaching phonetics... (I'm not sure about Sue)", where the focus would receive a so-called B-accent or rise-fall-rise accent (Jackendoff 1972, Büring 2003, Lai 2012, Wagner 2012a), could be represented as $\lambda\phi.\ commit(speaker, \phi)(teach(bill, phonetics))$, a statement about what the speaker is willing to commit to. In that case, the partition requirement holds.

assignment, and not merely focus. The innovation of this approach lies not in these general principles, but rather in the use of game theory and signaling to provide a unifying explanation in (18-a) of what counts as important or informative for these purposes. The focus game provides a way of rationally arriving at what information is maximally important, and the alignment between the crucial information and the assignment of prosodic focus relies on a simple linking hypothesis: that only the important information needs to be prominent in order for the whole intended meaning to survive possible effects of noise.[4]

We will see in 3.2 how the focus game can be iterated throughout the structure of a sentence to produce contrastive foci which are not tied to a QUD. But first, we begin illustrating the mechanics of the focus game by working through a simple example of QUD-motivated focus assignment.

3.1 Question-answer pairs

We begin by applying our game mechanics to a simple example in which the limitations placed by the QUD suffice to explain the observed focus pattern. Consider the following simple question-answer exchange.

(19) There are two professors, Bill and Sue, one of which teaches phonetics each semester.

Q: Who is teaching phonetics?
A: Bill is teaching phonetics.

We want our model to generate narrow focus on the subject, and to rule out any other focus patterns.

The QUD in this context delimits the set of felicitous responses to be the two-member set, $\{teach(bill, phonetics), teach(sue, phonetics)\}$. Let this be Φ, such that the union of the QUD and $\{\#\}$ is equal to T, which is in turn equal to A. This satisfies the first of our definitional constraints: non-babbling types and actions represent intended and interpreted meanings, respectively, both selected from a shared pool of formulae of the same semantic type. Our assumptions about the nature of QUDs satisfy the second and third constraints: the set of possible meanings is contextually available, and the space of possible meaningful Sender types creates a partition over an appropriate set of semantic objects. The fourth constraint, that the prior probability of a non-babbling type is proportional to the salience of the type, is trivially satisfied if we assume the possible QUD answers

[4] It would have been possible to build this linking hypothesis directly into the model by correlating focus with probabilities of surviving noise. However, this is an unnecessary complication of the model, given that we are dealing with categorical normative behavior. For example, imagine a model where non-focused elements have an 88% chance of surviving noise, while focused elements have a 99% chance. Given that at the level of abstract phonology, elements are either focused or not, the most efficient and effective message would give the elements that are crucial for successful interpretation the 99% chance, and leave the non-critical material with the 88% chance, not wasting effort on focusing elements which could be inferred in any case. Thus, we are licensed to simplify the model into one that only calculates the critical material, and then posit a linking rule that correlates this critical status with the assignment of focus. One research question which may be better answered by models that explicitly represent probabilities is the question of the relative phonetic prominence of foci. This is an interesting question for a different time.

	a_{BP}	a_{SP}	$a_{\#}$	$m_{\emptyset}$	m_B	m_S	m_P	m_{BP}	m_{SP}
t_{BP}	1,1	0,0	0,0	✓*	✓*	✗	✓*	✓*	✗
t_{SP}	0,0	1,1	0,0	✓*	✗	✓*	✓*	✗	✓*
$t_{\#}$	0,0	0,0	1,1	✓	✓	✓	✓	✓	✓

Fig. 6 Question-answer focus

to be equally salient, and correspondingly, that the non-babbling types have equal priors. Finally, we will begin our equilibrium derivation with a naive Receiver who reasons from a default belief that the denotation of the Sender's message is a component meaning of the Sender's type.

To pick out the crucial information within the sentence "Bill is teaching phonetics" in (19), we derive a separating equilibrium for the focus game represented in Fig.6.

There are two possible types in addition to the omnipresent low-probability babbling type ($t_{\#}$): there is t_{BP}, a Sender who wants to convey 'Bill is teaching phonetics', and t_{SP}, a Sender who wants to convey 'Sue is teaching phonetics.' The Receiver can select either of those two propositions as interpretations of S's message—a_{BP} and a_{SP}, respectively. Coordination occurs when the propositions picked out by type and action are the same, i.e. for type/action pairs t_{BP}/a_{BP} and t_{SP}/a_{SP}. We consider the following messages, in addition to a null message.

m_B	"Bill"
m_S	"Sue"
m_P	"is teaching phonetics"
m_{BP}	"Bill is teaching phonetics"
m_{SP}	"Sue is teaching phonetics"

Recall that this game is meant to model the selection of linguistic material that is optimally efficient for conveying relevant information, and therefore, the messages should not be conceived of as standalone utterances (though they could be used as such in many contexts), but rather fragments conveying the maximally important information that should be highlighted for the hearer.

For this game we obtain a separating equilibrium, pairing each non-babbling type with its unique best message: $\langle t_{BP}, m_B, a_{BP}\rangle$ and $\langle t_{SP}, m_S, a_{SP}\rangle$. That means, for type t_{BP} the message "Bill" is best, and for type t_{SP} the message "Sue" is best. The Receiver will assume any other message to have been produced by a babbler, and thus the corresponding focus structures are infelicitous.

$$\begin{aligned}
&R_0 \Rightarrow \{m_{\emptyset}, m_P \to \{t_{BP}, t_{SP}\}, m_B, m_{BP} \to \{t_{BP}\}, m_S, m_{SP} \to \{t_{SP}\}\}\\
&S_1 \Rightarrow \{t_{BP} \to \{m_B\}, t_{SP} \to \{m_S\}, t_{\#} \to M\}\\
&R_2 \Rightarrow \{m_B \to \{t_{BP}\}, m_S \to \{t_{SP}\}, m_{\emptyset}, m_P, m_{BP}, m_{SP} \to \{t_{\#}\}\}\\
&S_3 \Rightarrow \{t_{BP} \to \{m_B\}, t_{SP} \to \{m_S\}, t_{\#} \to \{m_{\emptyset}, m_P, m_{BP}, m_{SP}\}\}\\
&\text{Eq.} \Rightarrow \{\langle t_{BP}, m_B, a_{BP}\rangle, \langle t_{SP}, m_S, a_{SP}\rangle, \langle t_{\#}, \{m_{\emptyset}, m_P, m_{BP}, m_{SP}\}, a_{\#}\rangle\}
\end{aligned} \tag{10}$$

Let's briefly walk through how this was derived.

1. For R_0, who considers only whether message denotations are component meanings of possible types, both the null message and the message "phonetics" are compatible with either type. The two messages containing "Bill" ("Bill" and

"Bill is teaching phonetics") are uniquely compatible with 'Bill is teaching phonetics' (t_{BP}), while the two messages containing "Sue" are uniquely compatible with 'Sue is teaching phonetics' (t_{SP}).
2. For S_1, if her type is t_{BP}, she knows "Bill" is the optimally efficient message and guarantees coordination with R_0, and if her type is t_{SP}, she knows "Sue" is optimal. If she is of the babbling type, her type cannot be guessed by R_0, therefore a message is chosen at random.
3. Knowing what messages can be sent by S_1 Senders of different types, R_2 assumes "Bill" to have been produced by t_{BP}, "Sue" to have been produced by t_{SP}, and any other message to have been produced by a babbler.
4. An S_3 babbler will avoid m_B ("Bill") and m_S ("Sue"), which will both lead R_2 to guess an incorrect type.
5. At this point, further iteration will not affect how optimal messages are produced or interpreted.

We now know which portion of the utterance "Bill is teaching phonetics" should be selected for prosodic prominence assignment—the subject "Bill" is the optimal message in the corresponding focus game, i.e. it is part of a separating equilibrium where the Sender intends to convey that Bill is teaching phonetics. Thus, the subject is narrowly focused, giving rise to the following focus structure and intonation pattern.

(20) a. [Bill]$_\mathrm{F}$ is teaching phonetics
b. BILL is teaching phonetics.

As per (18-a), the substring "Bill" of "Bill is teaching phonetics" is a focus because $\langle t_{BP}, m_B, a_{BP}\rangle$ is part of a separating equilibrium in a focus game in which m_B is equivalent to "Bill", and both t_{BP} and a_{BP} are equivalent to ⟦"Bill is teaching phonetics"⟧. The infelicitous patterns below are ruled out simply on the grounds that they cannot be produced by a rational type t_{BP} Sender in this focus game.

(21) a. #Bill is teaching [phonetics]$_\mathrm{F}$
b. #Bill is [teaching phonetics]$_\mathrm{F}$
c. #Bill is [teaching]$_\mathrm{F}$ phonetics
d. #[Bill is teaching phonetics]$_\mathrm{F}$

Intuitively, (21-a), (21-b) and (21-c) are ruled out because the crucial information, "Bill", which distinguishes 'Bill is teaching phonetics' from the alternative 'Sue is teaching phonetics', is not part of the focus, while (21-d) is ruled out because, although "Bill" is part of the focus, broad focus is unnecessary and uneconomical—the subject alone is sufficient to highlight what is critical for interpretation.

The mechanics for single wh-questions are quite simple, but a full model must also be able to account for focus in multiple wh-questions as below.

(22) Q: Who's teaching what this semester?
A: BILL is teaching PHONETICS

The model must apply such that a pair of messages, "Bill" and "phonetics", can be produced, and must predict that this pair is uniquely optimal among competing pairs of messages for type t_{BP}.[5] Moreover, if the answer "Bill is teaching phonetics" is a partial answer to the wh-question, e.g., followed up with "and Sue is teaching phonology", then we must incorporate some notion of a sorting key (Kuno 1982) or strategy (Büring 2003) for answering the multiple wh-question in order to account for the placement of rise-fall-rise pitch contours (so-called contrastive topics). The proposal of Wagner (2012a) is a good candidate approach to such an integration, in that it separates the placement of foci as markers of pitch accent, which we are concerned with here, from the identification of the exact contour—i.e., fall or rise-fall-rise—associated with each pitch accent. A mapping from that proposal onto this one is possible if we allow for questioning strategies in the sense of Büring (2003) to prompt multiple focus games, where the order in which the games are set up determines where the B-accent falls.

For example, imagine that we first play a game where the set of alternatives is a set of wh-questions, {'What is Bill teaching?', 'What is Sue teaching?'}. Following Krifka (2001) and others, these questions can be represented as simple predicates, $\lambda x.teach(bill, x)$ and $\lambda x.teach(sue, x)$, respectively, such that the component meaninghood assumption can stand as is. Label these two meaningful types $t_{BT?}$ and $t_{ST?}$, and consider narrow focus messages m_B "Bill", m_S "Sue" and $m_{T?}$ "(is) teaching (what)", the null message, and broad focus messages $m_{BT?}$ and $m_{ST?}$. The exact same mechanics from above derive the following equilibrium.

$$\text{Eq.} \Rightarrow \{\langle t_{BT?}, m_B, a_{BT?}\rangle, \langle t_{ST?}, m_S, a_{ST?}\rangle, \langle t_{\#}, \{m_{\emptyset}, m_{T?}, m_{BT?}, m_{ST?}\}, a_{\#}\rangle\} \tag{11}$$

So the optimal message "Bill" will lead to the interpretation 'What is Bill teaching?', which then feeds a second focus game, this time with the QUD {'Bill is teaching phonetics', 'Bill is teaching phonology'}, which again by the exact same game mechanics will yield the optimal message "phonetics", ultimately producing a sequence of messages $\langle m_B, m_P\rangle$, yielding (22). The B-accent, then, will fall on whichever message is first in the sequence. A proper theory of B-accent placement is beyond our scope, but if we can separate accent placement from accent type, there is promise that the signaling model can account for multiple wh-questions as well.

In the following section we extend the focus game to the sub-sentential domain in a way that ameliorates the problem of delineating the message space by placing constraints on the semantic type of possible message denotations. We show that, by incorporating salience-based priors, we can derive contrastive foci within subconstituents and words by iteratively playing this constrained version of the focus game, deriving at each smaller nested level of structure an optimal message for transmitting the semantic content associated with that level. This allows us to use the same game mechanics to explain farmer sentences, and the probabilistic nature of the framework provides ways to account for some more subtle variants, which are addressed in 3.3.

[5] Of course, pairwise messages can only manifest phonologically, and do not license elision like the single focus cases do, simply because syntax does not allow the generation of "*Bill phonetics". A reviewer points out that such constructions are licit when gapping is involved, e.g., "Sue teaches phonology, and Bill phonetics", a difference which we assume for now to be purely syntactic.

3.2 Iterating focus games

We have shown how restrictions placed by a QUD allow the focus game to derive narrow focus within a full-sentence answer. We can also apply the same game mechanics to the word and phrase levels, asking the following question: what is the most crucial material for transmitting the meaning of a sub-sentential syntactic phrase or lexical item, assuming that the material above that phrase/word will be successfully interpreted? For example, consider the following variant of Rooth's (1992) famous example.

(23) An American farmer punched a Canadian farmer.

In an out-of-the blue context, broad default accent is optimal according to (18). But within that broad default accent structure we can also ask questions like the following: what information would be crucial for the transmission of the NP *American farmer* if this NP alone were subject to noise? To illustrate the intuition behind this question, imagine the following possible partial transmissions.

(24) a. An [______] punched a Canadian farmer.
b. An [American ______] punched a Canadian farmer.
c. An [______ farmer] punched a Canadian farmer.

One of these is more likely to be correctly interpreted than the others. It is clear that (24-a) is essentially hopeless—the missing NP could be *American farmer*, *Italian watchmaker*, etc. And (24-c) is not much better—one can surmise that the missing material is some nationality, but cannot differentiate from among a myriad of possibilities. The most salient nationality in the context is *Canadian*, but the resulting message, even ignoring the a/an distinction, "A Canadian farmer punched a Canadian farmer", is pragmatically somewhat odd. (One would expect something like, "A Canadian farmer punched another Canadian farmer.")[6] This brings us to (24-b). In principle, the empty slot created by noise could be filled with any kind of profession, but the most salient appropriate element to fill the missing slot would be *farmer*. This salience alone does not make "An American farmer punched a Canadian farmer" an iron-clad default guess for the whole intended utterance. However, if the hearer knows that the speaker chose their prominences strategically, with additional prominence on "American" such that only "American" survives the noise within the NP, then the hearer can reason as follows: The speaker would have stressed a non-salient meaning like "watchmaker" equally, if such a meaning was intended, therefore, the more salient interpretation 'American farmer' must have been the speaker's intent.

In other words, if we view the focus game as the selection of prominent material *within a noun phrase*, we should arrive at an equilibrium in which the adjective *American* constitutes an optimal message. We can propose accordingly that its status as a phrase-internal optimal message licenses a shift in relative prosodic prominence to *American* within that phrase. To think about it another way, *farmer* does not need to bear particularly strong prominence because it is doubly represented in the context. That is, if only one instance of *farmer* is obscured by noise,

[6] The example becomes fine in a context where it is particularly surprising that a Canadian farmer would punch another Canadian farmer in virtue of their Canadian-ness. See the discussion of example (42) in Section 5 for an account of these cases.

	a_{aF}	a_{aW}	a_{cF}	a_{cW}	$a_{\#}$	$m_{\emptyset}$	m_a	m_c	m_F	m_W	m_{aF}	m_{aW}	m_{cF}	m_{cW}
t_{aF}	1,1	0,0	0,0	0,0	0,0	✓	✓*	✗	✓	✗	✓*	✗	✗	✗
t_{aW}	0,0	1,1	0,0	0,0	0,0	✓	✓	✗	✗	✓	✗	✓*	✗	✗
t_{cF}	0,0	0,0	1,1	0,0	0,0	✓*	✗	✓*	✓*	✗	✗	✗	✓*	✗
t_{cW}	0,0	0,0	0,0	1,1	0,0	✓	✗	✓	✗	✓*	✗	✗	✗	✓*
$t_{\#}$	0,0	0,0	0,0	0,0	1,1	✓	✓	✓	✓	✓	✓	✓	✓	✓

Fig. 7 Focus within an NP

the other instance will allow the intended meaning to be reconstructed via a focus game at the NP level. It is more important from this perspective for *American* and *Canadian* to bear strong prominence, since they are not reconstructible in the same way.

We begin our analysis of (23) by looking only at the NP *American farmer*. A focus game for this NP is represented in Fig.7. For simplicity we consider only four meaningful types, each corresponding to a possible denotation for the noise-obscured NP: t_{aF} for 'American farmer', t_{aW} for 'American watchmaker', t_{cF} for 'Canadian farmer', and t_{cW} for 'Canadian watchmaker.' We consider the following meaningful messages.

m_a	"American"
m_c	"Canadian"
m_F	"farmer"
m_W	"watchmaker"
m_{aF}	"American farmer"
m_{aW}	"American watchmaker"
m_{cF}	"Canadian farmer"
m_{cW}	"Canadian watchmaker"

The NP-external utterance mentions a 'Canadian farmer', whose component meanings can be broken down as follows.

$\lambda x.Canadian(x)\ \&\ farmer(x)$
$\lambda x.Canadian(x)$
$\lambda x.farmer(x)$
$\lambda x.Canadian(x)\ \&\ farmer(x)$

We must now consider the role of salience in determining Receiver bias. We assume for now that only the NP-external intrasentential context ('a(n) P punched a Canadian farmer') determines salience. As suggested above, we can define a graded notion of salience by first establishing that one of the possible types in the game is already discourse-salient, designating that salient type as the default, and then calculating the degree of similarity between that type and other types based on how many component meanings are shared. Let t^* be the type in T (for simplicity we'll assume there is only one such type) which is entailed by a salient subset of of the common ground á la Schwarzschild (1999).

$$\text{salience}(t) = \frac{|\{\phi \mid \phi \rightarrow_{comp} t\} \cap \{\phi \mid \phi \rightarrow_{comp} t^*\}|}{|\{\phi \mid \phi \rightarrow_{comp} t^*\}|} \tag{12}$$

For the current game we obtain the following, where, due to presence of *Canadian farmer* in the NP-external context, $t^* = t_{cF}$.

$$\begin{aligned} \text{salience}(t_{aF}) &= 0.33 \\ \text{salience}(t_{aW}) &= 0 \\ \text{salience}(t_{cF}) &= 1 \\ \text{salience}(t_{cW}) &= 0.33 \end{aligned} \tag{13}$$

Based on this, we obtain information about the probability distribution over types.

$$\delta(t_{aW}) < \delta(t_{aF}), \delta(t_{cW}) < \delta(t_{cF}) \tag{14}$$

This hierarchy of salience-based priors is represented in Fig.7 via the placement of asterisks in each message column with multiple compatible types indicating which compatible type is maximally probable. Applying IBR to the game in Fig.7, we obtain a separating equilibrium selecting the crucial material within the target NP given the NP-external sentential context.[7]

$$\begin{aligned} R_0 \Rightarrow\ & \{m_a, m_{aF} \to \{t_{aF}\}, m_\emptyset, m_c, m_F, m_{cF} \to \{t_{cF}\}, \\ & m_W, m_{cW} \to \{t_{cW}\}, m_{aW} \to \{t_{aW}\}\} \\ S_1 \Rightarrow\ & \{t_{aF} \to \{m_a\}, t_{aW} \to \{m_{aW}\}, t_{cF} \to \{m_\emptyset\}, t_{cW} \to \{m_W\}, t_\# \to M\} \\ R_2 \Rightarrow\ & \{m_a \to \{t_{aF}\}, m_\emptyset \to \{t_{cF}\}, m_W \to \{t_{cW}\}, m_{aW} \to \{t_{aW}\}, \\ & m_{aF}, m_{cF}, m_{cW}, m_c, m_f \to \{t_\#\}\} \\ S_3 \Rightarrow\ & \{t_{aF} \to \{m_a\}, t_{aW} \to \{m_{aW}\}, t_{cF} \to \{m_\emptyset\}, t_{cW} \to \{m_W\}, \\ & t_\# \to \{m_{aF}, m_{cF}, m_{cW}, m_c, m_f\}\} \\ \text{Eq.} \Rightarrow\ & \{\langle t_{aF}, m_a, a_{aF}\rangle, \langle t_{aW}, m_{aW}, a_{aW}\rangle, \langle t_{cF}, m_\emptyset, a_{cF}\rangle, \langle t_{cW}, m_W, a_{cW}\rangle, \\ & \langle t_\#, \{m_{aF}, m_{cF}, m_{cW}, m_c, m_F\}, a_\#\rangle\} \end{aligned} \tag{15}$$

This shows that for the single NP *American farmer* situated within (23), "American" is the winner of the focus game. Note that this assumes that the meaning 'Canadian farmer' is made salient via cataphoric means, i.e., by considering what comes after "American farmer" in the utterance to be part of the context. The availability of an alternative conception of salience where such cataphoric reference is not possible accounts for the optionality of focus on "American", but not of focus on "Canadian", in the target sentence. If "Canadian farmer" is not considered as an antecedent, then no focus can be assigned at the "American farmer" node.

We are not limited to the NP level; in fact, we can iterate similar focus games throughout the structure of (23). This allows us to extend the general idea behind sub-sentential focus games: just as the focus game can encode how sentences relate to the discourse context by using QUDs to constrain types, the focus game can also encode how constituents within a sentence relate to each other (and potentially to the discourse context as well).

[7] Note that for the maximally salient meaning, 'Canadian farmer', the null message is optimal. This reflects the fact that, outside of an emphatic context (see Section 5), no contrastive foci should be assigned, e.g., in the sentence, "As usual, a Canadian farmer punched another Canadian farmer." This does not mean that no phonetic prominence at all should be assigned within an NP whose denotation is salient: See Selkirk (2007) for a discussion of default prominence assignment in the absence of foci.

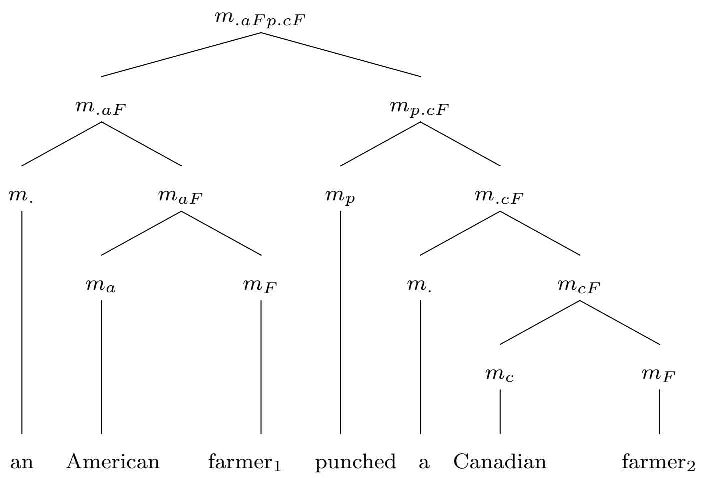

Fig. 8 Syntactic structure (distinct instances of *farmer* marked with subscripts)

We begin by establishing the basic syntactic constituency, shown in Fig.8. Each constituent is labeled with the above-used message notation, with a and c denoting 'American'/'Canadian', F denoting 'farmer', a dot denoting the meaning of the indefinite article *a*/*an*, and p denoting 'punched'. Using this notation, the utterance of the entire sentence, "an American farmer punched a Canadian farmer" is labeled $m_{.aFp.cF}$.

We iterate the game throughout the tree as follows.

– Beginning at the root node N of a sentence-level focus as determined by the question under discussion, assuming N has two daughters, A and B:
  1. If there exists a contextually available set Φ_N of mutually exclusive possible meanings of the same semantic type, where Φ_N contains $[\![N]\!]$:
     (a) Let M_N be a set of messages such that for all m in M_N, $[\![m]\!]$ is either a member of Φ_N, or else a component meaning of a member of Φ_N which is of the same semantic type as either $[\![A]\!]$ or $[\![B]\!]$.
     (b) Consider a focus game where $A = T = \Phi_N \cup \{\#\}$ and $M = M_N \cup \{m_\emptyset\}$; if a separating equilibrium exists, let WINNER(N) be the optimal message for type $t_{[\![N]\!]}$.
     (c) If either $[\![A]\!]$ or $[\![B]\!]$ is a component meaning of $[\![\textsc{winner}(N)]\!]$, then mark the corresponding daughter node as a winning node.
     (d) If either A or B are marked as winning nodes, repeat step 1. at the winning node(s), if they are branching.
  2. If no such contextually available set exists, repeat step 1. at any branching daughter nodes.
– After all iterations, any winning node that does not immediately dominate another winning node is marked as a focus.

The games are iterated such that only branching nodes and their immediate daughters are considered input to any individual focus game.[8] A QUD-based focus game

[8] The restriction that this places on M_N, along with anchoring the definition of component meaninghood to M_N, has the effect of considering only component meanings of the same

precedes the iteration, so that any nodes which are deemed informationally redundant given the QUD can be ignored. Focus is assigned only to winning nodes which don't immediately dominate other winning nodes in order to prevent unnecessary stacking of foci. Without this restriction, the iteration procedure would fail to take into account the informational overlap of various nodes, and would erroneously produce structures like (25-a), instead of the correct (25-b).

(25) Sue made some waffles, and...
  a. $[[\text{Bob}]_{\text{F}}$ $[\text{made}$ $[\text{some}$ $[\text{pancakes}]_{\text{F}}]_{\text{F}}]_{\text{F}}]$
  b. $[[\text{Bob}]_{\text{F}}$ made some $[\text{pancakes}]_{\text{F}}]$

For current purposes, let's assume a simplified context for (23) where there are two possible nationalities, Canadian and American, two possible professions, farmer and watchmaker, and two possible actions, punching and kicking. This means, for example, that at the VP node, the possible messages considered belong to one of three sets: (i) the set of possible VP denotations, each being of the form "punched/kicked a(n) American/Canadian farmer/watchmaker", (ii) the set of possible verbs, being simply "punched" vs. "kicked", and (iii) the set of possible direct objects, each being of the form "a(n) American/Canadian farmer/watchmaker".

First, we assume that in this context the QUD is a broad question like 'what happened?' and that accordingly, default prominence is assigned to the whole utterance. To determine whether any further foci exist further down the tree, we then begin the iteration at the root note $m_{.aFp.cF}$. We check whether there is a contextually available set of alternatives $\Phi_{aFp.cF}$ which contains the sentence meaning. Recall that we've defined "contextually available" as a property of a set of alternatives, satisfied if at least one member of the set is contextually salient. In the present context, this means there must be some antecedent denotation of the form, 'a(n) American/Canadian farmer/watchmaker punched/kicked a(n) American/Canadian farmer/watchmaker.' Assuming a neutral context, no such antecedent exists, and therefore a focus game cannot be constructed at this node.

We move on to $m_{p.cF}$. Similarly to above, there is no valid set of types that can be constructed, and thus we can skip down to $m_{.cF}$. Here, we can construct a valid set of types $T_{.cF} = \{t_{.aF}, t_{.cF}, t_{.aW}, t_{.cW}\}$, i.e. due to the salience of 'American farmer' in the node-external context, we have the contextually available set 'a(n)

semantic type as either N or one of its daughters. This not only makes the iteration manageable and simplifies the calculation of salience, but also has positive empirical consequences. Consider the following example from Schwarzschild (1999):

(i) Q: Who danced with John's mother?
  A: HER mother did.

The antecedent which activates the focus game at the DP node—"her mother"—is "John's mother", resulting in a 2x2 type space of the form $\{$'John's mother', 'John's father', 'John's mother's mother', $\cdots\}$. Because we are limited to considering possessives, NPs and combinations of the two, the fact that 'John's mother' is highly salient does not block the realization of focus on "her", because "her" is interpreted as the possessive 'John's mother's', which crucially is not salient given the restriction on M_N. Thus, for a type who wishes to convey 'John's mother's mother', focus falls straightforwardly on the possessor, and no focus is required on "mother", whose meaning can be recovered. Of course, 'John's mother' happens to be the most salient female in the global context, which licenses replacement of the possessive by the pro-form "her".

American/Canadian farmer/watchmaker' to choose from. The messages considered for this focus game are $m_{.cF}$ (= 'a Canadian farmer'), $m_{.}$ (= 'a'), m_{cF}, m_{aF}, m_{cW} and m_{aW}. Reflecting the fact that the indefinite article is common to all messages, we end up with the following separating equilibrium:

$$\begin{aligned}\text{Eq.} \Rightarrow \{&\langle t_{.aF}, m_{\emptyset}, a_{.aF}\rangle, \langle t_{.aW}, m_{aW}, a_{.aW}\rangle, \langle t_{.cF}, m_{cF}, a_{.cF}\rangle, \langle t_{.cW}, m_{cW}, a_{.cW}\rangle,\\ &\langle t_{\#}, \{m_{aF}, m_{.}\}, a_{\#}\rangle\}\end{aligned} \tag{16}$$

The winner at this node, then, is "Canadian farmer". By direct analogy, we can obtain the following equilibrium for $m_{.aF}$:

$$\begin{aligned}\text{Eq.} \Rightarrow \{&\langle t_{.aF}, m_{aF}, a_{.aF}\rangle, \langle t_{.aW}, m_{aW}, a_{.aW}\rangle, \langle t_{.cF}, m_{\emptyset}, a_{.cF}\rangle, \langle t_{.cW}, m_{cW}, a_{.cW}\rangle,\\ &\langle t_{\#}, \{m_{cF}, m_{.}\}, a_{\#}\rangle\}\end{aligned} \tag{17}$$

Finally, by analogy to m_{aF}, it is easy to calculate the separating equilibrium for the "Canadian farmer" node:

$$\begin{aligned}\text{Eq.} \Rightarrow \{&\langle t_{aF}, m_{\emptyset}, a_{aF}\rangle, \langle t_{aW}, m_{W}, a_{aW}\rangle, \langle t_{cF}, m_{c}, a_{cF}\rangle, \langle t_{cW}, m_{cW}, a_{cW}\rangle,\\ &\langle t_{\#}, \{m_{aF}, m_{aW}, m_{cF}, m_{a}, m_{F}\}, a_{\#}\rangle\}\end{aligned} \tag{18}$$

What we are left with in the end is the following set of winners.

$$\begin{aligned} t_{.aF} &\Rightarrow \text{“American farmer}_1\text{”}\\ t_{aF} &\Rightarrow \text{“American”}\\ t_{.cF} &\Rightarrow \text{“Canadian farmer}_2\text{”}\\ t_{cF} &\Rightarrow \text{“Canadian”}\end{aligned} \tag{19}$$

We ignore the winners that immediately dominate other winners, which yields the following focus structure.

(26) [An [American]$_{\mathrm{F}}$ farmer punched a [Canadian]$_{\mathrm{F}}$ farmer]

Again, note that we can account for the optionality of the contrastive focus on "American" if we limit antecedents to elements that have come *before* the element in question in the linear order of the sentence. Intuitively, salient antecedents can be determined globally, allowing for cataphoric instances such as "American farmer" above, where the salience of "farmer" anticipates the upcoming instance of "Canadian farmer", or salient antecedents can be determined locally, allowing only anaphoric instances, such that "Canadian farmer" is off limits because at the time of uttering "American farmer", it still has not been uttered yet. Rightly, this predicts that "Canadian" must be contrastively accented, no matter what.

All of this gives us a principled game-theoretic method of assigning abstract focus structures. It says nothing of the details of default prosody, nor of prosodic phrasing, interactions between prosody and syntactic structure, or word stress, all of which are important to the complete picture. A full and proper phonological treatment of the relevant examples is beyond this paper's scope, but we can indeed say something about prominence below the word level. Recall (7), reproduced below as (27).

(27) We drove from MINNE-sota to SARA-sota.

Here, the relative prominence of whole words is not the interesting desideratum for analysis. Rather, we want to derive the fact that the two contrasting place names, Minnesota and Sarasota, bear non-default stress at the lexical level. The intuition behind why this is the case is the same as for farmer sentences: the double representation of the phonological material /so.ta/ lessens the importance of accenting that material vis-à-vis attenuating the effect of noise in communication. And the same semantic constraints apply. For example, the following is not a valid use of contrastive focus, even though there is doubly represented phonological material, because 'Minnesota' and 'visit S.O.T.A.' are not possible alternatives.

(28) #We drove all the way from MINNE-sota just to VISIT s.o.t.a.
(s.o.t.a. = The Society of Typographic Aficionados)

This suggests that (27) should be derived using the same mechanics as used above. But to do so we must extend our assumptions slightly. Namely, we must allow the Receiver's default belief about the Sender's message to take one of two forms. Either (i) the message has a component meaning of the Sender's type (as heretofore assumed), or (ii) the message is a *component metrical foot* of a word signaling the Sender's type. (see Artstein 2004, on why metrical feet are the right unit of analysis). This allows us to iterate the focus game even further, below the syntactic level into the word level. There, we can derive, for example, that Minne- is a better foot to stress than -sota given the context in which Sarasota is salient. The following equilibrium is calculated by analogy to the equilibrium calculation for Fig.7, and assumes four possible types, each denoting a place name within the United States: Minnesota, Sarasota, Minnetoga and Saratoga.

$$
\begin{aligned}
R_0 \Rightarrow \{ & m_{Minne}, m_{Minnesota} \rightarrow \{t_{Minnesota}\}, \\
& m_{\emptyset}, m_{Sara}, m_{sota}, m_{Sarasota} \rightarrow \{t_{Sarasota}\}, \\
& m_{toga}, m_{Saratoga} \rightarrow \{t_{Saratoga}\}, \\
& m_{Minnetoga} \rightarrow \{t_{Minnetoga}\}\} \\
S_1 \Rightarrow \{ & t_{Minnesota} \rightarrow \{m_{Minne}\}, t_{Minnetoga} \rightarrow \{m_{Minnetoga}\}, \\
& t_{Sarasota} \rightarrow \{m_{\emptyset}\}, t_{Saratoga} \rightarrow \{m_{toga}\}, t_{\#} \rightarrow M\} \\
R_2 \Rightarrow \{ & m_{Minne} \rightarrow \{t_{Minnesota}\}, m_{\emptyset} \rightarrow \{t_{Sarasota}\}, \\
& m_{toga} \rightarrow \{t_{Saratoga}\}, m_{Minnetoga} \rightarrow \{t_{Minnetoga}\}, \\
& m_{Sarasota}, m_{Minnesota}, m_{Saratoga}, m_{Sara}, m_{sota} \rightarrow \{t_{\#}\}\} \\
S_3 \Rightarrow \{ & t_{Minnesota} \rightarrow \{m_{Minne}\}, t_{Minnetoga} \rightarrow \{m_{Minnetoga}\}, \\
& t_{Sarasota} \rightarrow \{m_{\emptyset}\}, t_{Saratoga} \rightarrow \{m_{toga}\}, \\
& t_{\#} \rightarrow \{m_{Sarasota}, m_{Minnesota}, m_{Saratoga}, m_{Sara}, m_{sota}\}\} \\
\text{Eq.} \Rightarrow \{ & \langle t_{Minnesota}, m_{Minne}, a_{Minnesota}\rangle, \\
& \langle t_{Minnetoga}, m_{Minnetoga}, a_{Minnetoga}\rangle, \\
& \langle t_{Sarasota}, m_{\emptyset}, a_{Sarasota}\rangle, \langle t_{Saratoga}, m_{toga}, a_{Saratoga}\rangle, \\
& \langle t_{\#}, \{m_{Sarasota}, m_{Minnesota}, m_{Saratoga}, m_{Sara}, m_{sota}\}, a_{\#}\rangle\}
\end{aligned} \tag{20}
$$

Similarly, the focus game for the context in which Minnesota is salient will produce an equilibrium where Sara- is the best foot to stress in "Sarasota". If we iterate through (27), assuming a rather vanilla, out-of-the-blue utterance context, this time allowing for focus games at the level of individual lexical items which dominate metrical feet, we can obtain the following.

(29) We drove from {[Minne]$_{\mathrm{F}}$ sota} to {[Sara]$_{\mathrm{F}}$ sota}

Since focus requires elevated relative prominence, we must violate the default stress pattern associated with the lexical items "Minnesota" and "Sarasota".

By iterating the focus game below the sentence and word levels we are able to make predictions about intonation beyond the question-answer pairs and exhaustive focus examples covered by Schmitz (2008) and Bergen & Goodman (2014), further supporting the usefulness of a strategic approach to explaining the interaction between pragmatics and phonology.

3.3 Competing antecedents

Our model assumes that spaces of possible types/actions in the focus game must be "activated" in the discourse context, and that typically this is accomplished via some antecedent. We have used the existence of an activating antecedent to anchor an operationalization of a gradient notion of salience. Because salience is gradient rather than categorical, a distinguishing property of this approach over e.g. Schwarzschild (1999), we predict situations with multiple competing antecedents where one antecedent "wins out" over the others in terms of its effect on focus placement, in virtue of its degree of salience being higher. Indeed we do find such cases, as in (30).

(30) An [American]$_{\mathrm{F}}$ farmer punched a [Canadian]$_{\mathrm{F}}$ farmer, and then...
a. an American [watchmaker]$_{\mathrm{F}}$ punched a Canadian [watchmaker]$_{\mathrm{F}}$
b. ??an [American]$_{\mathrm{F}}$ watchmaker punched a [Canadian]$_{\mathrm{F}}$ watchmaker

At issue here is the somewhat counter-intuitive fact that (30-a), in which both instances of "watchmaker" are in focus, is more natural than (30-b). To illustrate why this is, consider the "Canadian watchmaker" node. There are three suitable antecedents: 'American farmer', 'Canadian farmer' and 'American watchmaker'. If we limited ourselves to the antecedent which occurs in the same sentence ('American watchmaker'), then we would wrongly predict (30-b) to be optimal, analogous to our derivation of the simple farmer sentence. If we were instead to simply average the salience with respect to all three antecedents, we would obtain the following:

$$\begin{aligned} \text{salience}(t_{aF}) &= 0.55 \\ \text{salience}(t_{aW}) &= 0.44 \\ \text{salience}(t_{cF}) &= 0.44 \\ \text{salience}(t_{cW}) &= 0.22 \end{aligned} \tag{21}$$

This also makes the wrong prediction: Focus on "watchmaker" should suggest an interpretation of "American watchmaker", and thus the optimal message at the "Canadian watchmaker" node should be the entire NP.

But there is no reason to think that the δ function, which encodes a prior probability distribution over possible types, should depend on simple averaging.

A much more sensible approach is to take the weighted average of the different salience values, weighted by how salient the antecedents themselves are. This requires us to determine the conditions on higher vs. lower relative salience, and one obvious candidate source of such variation is structural parallelism. That is, when playing the focus game at a direct object node, an antecedent in a direct object position should be given greater weight than other antecedents. This is the same sort of principle at work in pronoun resolution, as in the following, where it is preferred that the pronoun in subject position refer back to the subject antecedent.

(31) John smiled at Bill, and then he insulted him.

For illustration's sake, let's imagine that the direct object antecedent "Canadian farmer" is given four times as much weight as the other two. We end up with the following salience values at the "Canadian watchmaker" node.

$$\begin{aligned} \text{salience}(t_{aF}) &= 0.44 \\ \text{salience}(t_{aW}) &= 0.22 \\ \text{salience}(t_{cF}) &= 0.72 \\ \text{salience}(t_{cW}) &= 0.28 \end{aligned} \tag{22}$$

Now it is clear that "watchmaker" leads to the interpretation 'Canadian watchmaker' at this node, as more weight has been given to "Canadian farmer" as an antecedent. Applying this principle to the whole sentence in this context will predict (30-a). Note that if we were to switch the subject and object in (30-a), depending on the weight given to structurally parallel antecedents, we would predict foci on both the nationality and the profession. Intuitions about these cases can get somewhat subtle,[9] but we hold that this multiple focus structure indeed becomes more natural when the subject and object are switched.

(32) An [American]$_{\text{F}}$ farmer punched a [Canadian]$_{\text{F}}$ farmer, and then...
  a. A [Canadian]$_{\text{F}}$ [watchmaker]$_{\text{F}}$ punched an [American]$_{\text{F}}$ [watchmaker]$_{\text{F}}$
  b. ??A Canadian [watchmaker]$_{\text{F}}$ punched an American [watchmaker]$_{\text{F}}$

Without allowing for higher and lower degrees of salience, it is unclear how to account for these otherwise unexpected stresses on *watchmaker* in any case. This is one area where, beyond its higher-level explanatory advantages, the signaling model offers potential empirical advantages as well.

Now, having outlined the core of the iterated focus game model, we turn to matters of focus association and second occurrence focus.

[9] For example, (32-a) below is most naturally pronounced with a tapering off of pitch accent strength at the end; however, neither "American" nor "watchmaker" are completely deaccented, and the slight decrease in prominence can be seen as a reflex of given status. Thus, the standard analysis would claim that these are proper foci. (See the discussion of second occurrence focus in Section 4.)

4 Association and second occurrence focus

4.1 Association with *only*

In the previous section we proposed focus games as an interactive model of strategic focus placement in answers to questions and utterances containing semantically contrasting elements. We now address a particular phenomenon often evoked by discussions of this kind: the apparent conventional association of lexical items such as *only* with prosodic foci (Rooth 1992, Beaver & Clark 2008, and many others). The following illustrates.

(33) a. The governor only gave contracts to the MAFIA.
b. The governor only gave CONTRACTS to the Mafia.

It is clear that in (33-a), the intended meaning is that the governor gave contracts to the Mafia, and to no one else (and therefore was probably charged with corruption). In (33-b), on the other hand, the intended meaning is that the governor gave contracts to the Mafia, and gave nothing more than that to the Mafia (implying that they wanted more from the governor). Following Beaver & Clark (2008), we may posit that the meaning of 'only ϕ' is something like, 'ϕ is true, as opposed to a more informative alternative', where the notion of "informative" can be thought of as measuring the number of entailments, such that 'the governor gave contracts to everyone' is more informative than 'the governor gave contracts to the Mafia'.

Under such an account, the interpretation of *only* relies on the contextual availability of some set of alternative propositions which are ordered by informativity. For any such set, a valid space of possible focus game types/actions can be constructed by converting the set of propositions into a partition over worlds, such that each proposition entails the falsity of the other. Thus, the meaning of 'only ϕ' can be re-stated in focus game terms. We tentatively propose that 'only' combines with some proposition ϕ to trigger an exhaustive focus game for a contextually available set of possible Sender types T, where T contains ϕ. For (33-a), this means something like the following.

(34) $[\![(33\text{-a})]\!] \Rightarrow$ 'The governor gave contracts to the Mafia' is a true proposition in the contextually available set of possible Sender types $T = \{$'The governor gave contracts to the Mafia (and no one else)', 'The governor gave contracts to everyone'$\}$

If stated in these terms, then the correlation of focus is unsurprising, because different values for T will induce different focus games, and therefore different prosodic outputs. But the relationship is indirect: there is no encoding of focus marking directly in the meaning of *only*, but rather an implication that the state of the discourse is such that a certain kind of focus game can be played at the level at which *only* adjoins. By hypothesis, if a focus game can be played, then a particular prosodic pattern emerges.

To illustrate how this might work game-theoretically, consider a variant of our focus game in which there are two decisions for the hearer to make, i.e., two aspects to S's type which are unknown to R: (i) the semantic content of S's type t, and (ii) the set T from which t was drawn, i.e., the association of *only*. Assume an applicative structure like the following for (33), and consider what happens at the ApplP node.

(35) The governor only [$_{\text{VP}}$ gave [$_{\text{ApplP}}$ contracts to the Mafia]]

Let's consider a set of types such that the possible ApplP arguments are 'contracts' and 'indictments' for the patient and 'Mafia' and 'police' for the recipient, where, for example $t_{C\underline{M}}$ denotes the meaning 'contracts to the mafia' selected from the type set $T = \{$'contracts to the mafia', 'contracts to the police'$\}$:

$t_{C\underline{M}}$	'contracts to the mafia (not the police)'
$t_{\underline{C}M}$	'contracts (and not indictments) to the mafia'
$t_{I\underline{M}}$	'indictments to the mafia (not the police)'
$t_{\underline{I}M}$	'indictments (and not contracts) to the mafia'
$t_{C\underline{P}}$	'contracts to the police (not the mafia)'
$t_{\underline{C}P}$	'contracts (and not indictments) to the police'
$t_{I\underline{P}}$	'indictments to the police (not the mafia)'
$t_{\underline{I}P}$	'indictments (and not contracts) to the police'

Messages can be one or both of the ApplP arguments, ignoring the obligatory preposition, creating $M = \{m_C, m_M, m_{CM}, m_I, m_{IM}, m_P, m_{CP}, m_{IP}\}$.

The addition of different possible associations of *only* requires us to enrich our notion of message-type compatibility for purposes of specifying R_0. First, as before, we assume that message denotations are component meanings of types. But we must also make an assumption about how the naive hearer will map S's message onto association possibilities. We have conceived of associations as type sets represented by S, and R_0 can use these type set representations to reason about what types could send which messages. For example, type $t_{C\underline{M}}$ draws from $T' = \{$'contracts to the mafia', 'contracts to the police'$\}$, while type $t_{\underline{C}M}$ draws from $T'' = \{$'contracts to the mafia', 'indictments to the mafia'$\}$. Analogously, type $t_{I\underline{M}}$ draws from $T''' = \{$'indictments to the mafia', 'indictments to the police'$\}$, while type $t_{\underline{C}P}$ draws from $T'''' = \{$'contracts to the police', 'indictments to the police'$\}$. R_0 can play 'mini-games' on each one of those sets to determine what the optimal message would be for a given type given its association set. This means that message m_I would be assumed to have come from either $t_{\underline{I}M}$ or $t_{\underline{I}P}$, message m_M would be assumed to have come from either $t_{I\underline{M}}$ or $t_{C\underline{M}}$, etc. Putting it together, we can specify R_0 as follows.

$$\begin{aligned} R_0 \Rightarrow \{ & m_{IM} \to \{t_{I\underline{M}}, t_{\underline{I}M}\}, m_I \to \{t_{\underline{I}M}, t_{\underline{I}P}\}, m_M \to \{t_{I\underline{M}}, t_{C\underline{M}}\}, \\ & m_{CM} \to \{t_{C\underline{M}}, t_{\underline{C}M}\}, m_C \to \{t_{\underline{C}M}, t_{\underline{C}P}\}, m_{IP} \to \{t_{I\underline{P}}, t_{\underline{I}P}\}, \\ & m_P \to \{t_{I\underline{P}}, t_{C\underline{P}}\}, m_{CP} \to \{t_{C\underline{P}}, t_{\underline{C}P}\}, m_\emptyset \to T \setminus \{t_\#\}\} \end{aligned} \tag{23}$$

S's best response to this, and R's best response to S's best response, are given below.

$$\begin{aligned} S_1 \Rightarrow \{ & t_{C\underline{M}}, t_{I\underline{M}} \to \{m_M\}, t_{\underline{C}M}, t_{\underline{C}P} \to \{m_C\}, \\ & t_{\underline{I}M}, t_{\underline{I}P} \to \{m_I\}, t_{C\underline{P}}, t_{I\underline{P}} \to \{m_P\}, t_\# \to M\} \\ R_2 \Rightarrow \{ & m_M \to \{t_{C\underline{M}}, t_{I\underline{M}}\}, m_C \to \{t_{\underline{C}M}, t_{\underline{C}P}\}, m_I \to \{t_{\underline{I}M}, t_{\underline{I}P}\}, m_P \to \{t_{C\underline{P}}, t_{I\underline{P}}\}, \\ & m_{CM}, m_{CP}, m_{IM}, m_{IP}, m_\emptyset \to \{t_\#\}\} \end{aligned} \tag{24}$$

Another round of iteration will get to the following equilibrium. It is not a separating equilibrium, as there is an unresolvable ambiguity about S's type given most messages. However, there is still a unique optimal message for each non-babbling type to send.

$$\begin{aligned}\text{Eq.} \Rightarrow \{&\langle t_{C\underline{M}}, m_M, \{a_{C\underline{M}}, a_{I\underline{M}}\}\rangle, \langle t_{\underline{C}M}, m_C, \{a_{\underline{C}M}, a_{\underline{C}P}\}\rangle,\\ &\langle t_{I\underline{M}}, m_M, \{a_{C\underline{M}}, a_{I\underline{M}}\}\rangle, \langle t_{\underline{I}M}, m_I, \{a_{\underline{I}M}, a_{\underline{I}P}\}\rangle,\\ &\langle t_{C\underline{P}}, m_P, \{a_{C\underline{P}}, a_{I\underline{P}}\}\rangle, \langle t_{\underline{C}P}, m_C, \{a_{\underline{C}M}, a_{\underline{C}P}\}\rangle,\\ &\langle t_{I\underline{P}}, m_P, \{a_{C\underline{P}}, a_{I\underline{P}}\}\rangle, \langle t_{\underline{I}P}, m_I, \{a_{\underline{I}M}, a_{\underline{I}P}\}\rangle,\\ &\langle t_{\#}, \{m_{CM}, m_{CP}, m_{IM}, m_{IP}, m_{\emptyset}\}, a_{\#}\rangle\}\end{aligned} \tag{25}$$

What this tells us is that, from the standpoint of sending strategically optimal messages in a signaling game, the best the Sender can do is to focus the element which is associated with *only*, because any additional foci would only create extra effort without increasing the probability of successful coordination.[10] Of course, the situation is not so dire in most real communicative situations, in that as the true effect of noise is rather small—this is the topic we turn to in Section 5. But first, we show how this association-representing variant of the focus game yields some nice empirical results regarding second occurrence focus.

4.2 Second occurrence focus

One *prima facie* challenge to the focus game approach is the existence of so-called second occurrence foci. Consider the following.

(36) The governor only gave contracts to the MAFIA…
 a. whereas the MAYOR only gave INDICTMENTS to them.
 b. #whereas the MAYOR only gave INDICTMENTS to 'em. ("them" = the Mafia and no one else)

The salience of the previously mentioned Mafia allows an unaccented pronoun to be used in (36-a), even though the interpretation is such that 'the Mafia' is being exhaustively and contrastively identified as the recipients of the mayor's indictments. This alone should pose no problem; the current approach predicts no necessary accent on *them*, because 'the Mafia', being salient, should be reconstructible in the same way that 'farmer' and /so.ta/ are in previous examples. However, (36-b) shows that, in fact, some degree of prominence is required on the pronoun, such that it cannot be phonetically reduced to "'em" without changing the interpretation.

Can our strategic model of focus placement account for this? To see how it might, let's consider the above variant of the focus game where there is a two-way ambiguity with respect to both the core content conveyed by S's type as well as which element is associated with *only*.

The only difference between the game in 4.1 and the game we construct to model a second occurrence focus context is a less uniform δ function. Just as with the farmer sentences in 3.2, we introduce a δ ranking for meaningful types based on salience from the utterance context. In this case, δ is higher for types with recipient association and for which 'contracts' and 'Mafia' are component meanings, such

[10] For simplicity, we have ignored the possibility of a long but completely unambiguous message like, "contracts to the mafia and to no one else".

that $\delta(t_{C\underline{M}})$ is maximal and $\delta(t_{\underline{I}P})$ is minimal. As above, the model does not converge to a clear separating equilibrium, but it nonetheless gives us a stable set of optimal messages for each type.

$$
\begin{aligned}
R_0 \Rightarrow \{ & m_{IM} \to \{t_{I\underline{M}}\}, m_I \to \{t_{\underline{I}M}\}, m_M, m_{CM} \to \{t_{C\underline{M}}\}, m_C \to \{t_{\underline{C}M}\}, \\
& m_{IP} \to \{t_{I\underline{P}}\}, m_P, m_{CP} \to \{t_{C\underline{P}}\}, m_\emptyset \to \{t_{C\underline{M}}\}\} \\
S_1 \Rightarrow \{ & t_{I\underline{M}} \to \{m_{IM}\}, t_{\underline{I}M} \to \{m_I\}, t_{C\underline{M}} \to \{m_\emptyset\}, t_{\underline{C}M} \to \{m_C\}, \\
& t_{I\underline{P}} \to \{m_{IP}\}, t_{\underline{I}P} \to \{m_I, m_P, m_{IP}\}, t_{C\underline{P}} \to \{m_P\}, t_{\underline{C}P} \to \{m_C, m_P, m_{CP}\}, \\
& t_{\#} \to M\} \\
R_2 \Rightarrow \{ & m_{IM} \to \{t_{I\underline{M}}\}, m_I \to \{t_{\underline{I}M}\}, m_M, m_{CM} \to \{t_{\#}\}, \\
& m_C \to \{t_{\underline{C}M}\}, m_{IP} \to \{t_{I\underline{P}}\}, m_P \to \{t_{C\underline{P}}\}, m_{CP} \to \{t_{\underline{C}P}\}, m_\emptyset \to \{t_{C\underline{M}}\}\} \\
S_3 \Rightarrow \{ & t_{I\underline{M}} \to \{m_{IM}\}, t_{\underline{I}M} \to \{m_I\}, t_{C\underline{M}} \to \{m_\emptyset\}, t_{\underline{C}M} \to \{m_C\}, \\
& t_{I\underline{P}} \to \{m_{IP}\}, t_{\underline{I}P} \to \{m_I, m_P, m_{IP}\}, t_{C\underline{P}} \to \{m_P\}, t_{\underline{C}P} \to \{m_{CP}\}, \\
& t_{\#} \to \{m_M, m_{CM}\}\}
\end{aligned}
\tag{26}
$$

This makes some predictions which at first seem rather odd, but which turn out to be empirically supported. The first is that type $t_{\underline{I}P}$ cannot succeed in this game. We assume in the above derivation that all things being equal, $t_{\underline{I}P}$ randomly selects a message whose denotation is a component meaning of $t_{\underline{I}P}$, however, the Receiver will never successfully interpret any of those messages. A more realistic scaled up model which allows for circumlocutions in these instances would predict, quite plainly, that the string "the mayor only gave indictments to the police"—no matter what the focus structure is—cannot be used to convey the meaning 'the mayor gave indictments (and only indictments) to the police' in this second occurrence context. This seems to be true.

(37) The governor only gave contracts to the MAFIA, and the MAYOR...
 a. #only gave INDICTMENTS to the police ($t_{\underline{I}M}$)
 b. #only gave indictments to the POLICE ($t_{C\underline{P}}$)
 c. only gave INDICTMENTS to the POLICE ($t_{I\underline{P}}$)

The first two possibilities lead to the assumption of a semantically wrong type, while the third possibility can only be interpreted such that *only* associates with *police*.

Relatedly, this means that analogous focus patterns in second occurrence contexts can lead to different associations depending on whether one of the focused elements is salient. This is shown below.

(38) The governor only gave contracts to the MAFIA, and the MAYOR...
 a. only gave INDICTMENTS to the POLICE ($t_{I\underline{P}}$)
 b. only gave CONTRACTS to the POLICE ($t_{\underline{C}P}$)

In other words, the signaling model predicts that "CONTRACTS to the POLICE" will signal that the role associated with *only* has shifted from the associated role in the antecedent. Admittedly this is not the most natural of utterances, but with a little bit of context, we can get this shift in association quite naturally.

(39) The city was mired in controversy last year after high-ranking officials accepted payoffs for giving contracts and holding back indictments of suspected lawbreakers. For example... The governor only gave contracts to the [Mafia]$_{\mathrm{F}}$, while the mayor only gave [contracts]$_{\mathrm{F}}$ to the [police]$_{\mathrm{F}}$

Crucially, we see that the requirement of two foci in (36-a) is correctly predicted. It is necessary in the second-occurrence context to focus both 'indictments' and 'mafia' in order to avoid mistakenly signaling an interpretation where *only* associates with patients and not recipients. However, the focus (on 'Mafia') which is discourse salient is phonetically dampened to avoid placing too much phonetic prominence on a given element (see Selkirk 2007)).

5 Inferences from off-equilibrium messages in low-noise environments

The focus game tells us, for a given context and a given intended meaning, what the most important or noteworthy information is for purposes of recovering the intended meaning from a minimal signal using the context. By convention, this is what should be focused, and the obvious intuition behind this, going back to Claude Shannon and other work in information theory, is that this helps ameliorate the effect of noise, to the extent that phonological focus correlates (albeit imperfectly) with phonetic prominence (recall Fig.1 from 1.1). By hypothesis, these conventions, while ultimately rooted in a normative game-theoretic model, are employed somewhat automatically in natural language usage—it would be a wild claim that speakers are running an IBR model in their heads whenever they use contrastive focus in speech. Perhaps this is obvious, but it is worth highlighting by way of pointing out a crucial fact about focus: although the optimality of prosodic focus marking in languages whose grammars allow it is rooted in considerations of noise, focus is nonetheless mandatorily marked even in *low-noise environments*. This has a variety of interesting consequences for how focus can be used and interpreted. A low-noise environment is a context in which the Sender believes, for whatever reasons, that his/her utterance will be fully received by the Receiver. What we want to add to our model is a mechanism which: (i) enforces the selective marking of focus even in these environments, but (ii) allows the Sender to exploit his/her confidence in successful transmission to convey non-literal meanings.

We have assumed that a categorical evaluation of focus structures is possible, i.e., if a type t Sender would not send F in a focus game G, then F is a bad focus structure for an utterance that gives rise to G. In this section, we consider a larger three-step model of focus selection, with a more conservative criterion for predicting an utterance to be infelicitous in virtue of its focus structure:

1. First, a focus game G is played as before.
2. Then, with some probability which is inversely proportional to the level of noise, the Sender's true type is *revealed* to the Receiver.
3. If $\langle t, m \rangle$ does not reflect an equilibrium strategy for an S of type t, then R searches for some other, minimally different game G' that S might have been playing instead.
   – If no such game is found, then the utterance at large is infelicitous.
   – If a suitable G' is found, R plays G' instead.

This larger model was not necessary for the simpler examples in the above sections, because we were specifically dealing with contexts with no obvious candidate for a G' to explain possible off-equilibrium messages. That changes when we consider the following examples.

(40) An American farmer was talking to an INTELLIGENT farmer

(41) John called Mary a Republican, then SHE insulted HIM

(42) a. We saw AMERICANS betraying AMERICANS!
b. #We saw AMERICANS saying hi to AMERICANS

Consider the ways in which a "minimally different" game G' can be constructed.

1. One could adjust the space of possible types and actions.
2. One could adjust the content of the surrounding discourse context
3. One could adjust the terms of the δ function.

Let's address these in turn. First, example (40) illustrates an inference that arises from an alteration of the space of possible types and actions. As per the partition requirement on well-formed spaces of possible types/actions, the literal meanings of *American farmer* and *intelligent farmer* cannot belong to the same T/A, because they are semantically orthogonal. So when a focus game is played, for example, at the object NP node, resulting in a contrastively focused *intelligent*, the hearer is forced to coerce T/A into a proper partition. But without the speaker's true type being revealed, there is ambiguity as to how this can be done. Consider two equally simple ways: first, the hearer could partition the space of occupations into $\lambda x.\ farmer(x)$ & $unintelligent(x)$ and a number of predicates of the form $\lambda x.\ P(x)$ & $intelligent(x)$, where P is an occupation. In other words, the partition could be between farmers who are inherently unintelligent, and other occupational groups who are inherently (more) intelligent, i.e., the inference that would be drawn from the utterance, "an American farmer was talking to an INTELLIGENT American." Alternatively, the hearer could partition the space of nationalities into $\lambda x.\ American(x)$ & $unintelligent(x)$ and $\lambda x.\ P(x)$ & $intelligent(x)$, where P is a nationality: a partition between Americans, who are inherently unintelligent, and citizens of other nations who are inherently (more) intelligent. If the speaker's true type has been revealed, then it becomes clear that only the second partition is possible, forcing the inference that (contrary to the authors' opinion) Americans tend to lack intelligence. The persistence of a strategy for focus selection rooted in focus games even in low-noise environments allows a speaker who is confident of his/her true type being revealed regardless of message to exploit the structure of the game to convey non-literal meaning.

Example (41) has a similar flavor, in that it requires a coercion of a type space into a partition (in this case at the VP node, *insulted him*), but with two differences. First, the identity of the partition itself, something like $\{\lambda x.\ insult(x, mary)$, $\lambda x.\ praise(x, mary)$, $\lambda x.\ insult(x, john)$, $\lambda x.\ praise(x, john)\}$, is completely unclear until the Sender's true type has been revealed, since the crucial verbal material bears no focus. Second, it requires a coercion of the context, in this case into one which makes 'insulting' salient, precluding its inclusion in the focus. This is accomplished via a pragmatic enrichment of 'called Mary a Republican', whereby that predicate is taken to entail an insult to Mary. The hearer knows (that the speaker knows that the hearer knows) that this pragmatic enrichment is intended, leading to the inference that calling someone a Republican is, by its nature, insulting.

Finally, example (42) illustrates the possibility of drawing inferences by expanding the notion of salience encoded in the δ function. Intuitively, the license for the contrastive stress on *Americans* in (42-a) is an element of surprise regarding the semantic contribution of the focused arguments: an American would never betray one of their own, thus ‘American’ is a surprising betrayer/betrayee of an American. In other words, the speaker is implying that the NP slots represented in (42) should not have been expected to be filled by the meaning ‘American’. This is a different notion of salience from what we have been using. In the canonical focus game, salience is assessed only with respect to the immediate linguistic context, but it is also possible to encode a similar notion that is assessed with respect to some general world knowledge. Under this latter conception, δ is very low for *Americans* when the interlocutors’ world knowledge contains few instances of Americans betraying each other. The banality of (42-b) accounts for the contrast with (42-a), in that no reasonable δ function can be constructed for which *Americans* is not already salient enough at either DP/NP node to preclude contrastive focus. It is important that this version of salience come into play only as a fallback, when the observed focus pattern cannot be an equilibrium message for the appropriate type in a focus game where only salience based on linguistic context is at play. If the linguistic context did not take precedence, then we would expect to find focus patterns like the following.

(43) ??An AMERICAN seller sold a CANADIAN buyer an antique vase.

If we defaulted to the world-knowledge conception of salience, δ would be higher for sellers at the subject node, and higher for buyers at the indirect object node, since we know that buyers sell things to sellers, and thus (43) should be acceptable. The problem is solved if we give precedence to linguistic salience. Consider the subject NP node’s focus game. The immediate surrounding linguistic context (‘A… sold a Canadian buyer an antique vase’) makes ‘buyer’, rather than ‘seller’ the most salient role in the transaction. Not taking into account world knowledge about buying/selling, “American” becomes the optimal message for a type conveying ‘American buyer’, and thus (43) would lead to a misinterpretation in any noisy environment which deletes “seller”.

In this case our model hearer will only consider world knowledge if the speaker’s true type is revealed, i.e., in a low-noise environment. And though a G' can be constructed with an expanded δ in which (43) becomes felicitous, crucially, the hearer learns no new information from the speaker’s invoking G'. There is no inference to be drawn from the choice of G', no good reason for not sticking with G. So (43) remains pragmatically quite odd. But is possible to exploit the expanded notion of δ to convey meanings of surprise as in (42), with a configuration similar to (43), as below.

(44) We knew something was amiss when an AMERICAN seller sold a SOVIET buyer an antique vase.

This is perhaps not the most natural intonation, but it seems better than (43), likely on the grounds that there is some additional meaning that can be uncovered: that the speaker finds the situation to be particularly surprising or suspicious.[11]

We end by noting that no claims are made about whether or how interlocutors assess whether they are in a low-noise or high-noise environment. As already discussed, the model presented here, like most game-theoretic models, is normative, a model of what ideal conversational agents would do, intended to shed light on why the system of prosodic focus placement developed to be the way that it is and not some radically different way. In real-time language use, the optimal strategies in these games are surely encoded in the minds of speakers and hearers not directly, but rather in the form of heuristics which employ these strategies imperfectly and automatically, without much overt reasoning. It is therefore not too surprising that our intuitions about the more complex cases begin to become fuzzier. Nonetheless, we have situated a variety of facts about focus placement within a general strategic framework, showing that a wide variety of related behaviors instantiate equilibrium strategies in a focus game.

6 Conclusion

In this paper we have expanded upon the role of noise in communication by modeling focus placement in intonation languages as the selection of crucial information via a signaling game. Using established game-theoretic techniques we have argued for a mapping between equilibrium strategies in these games and observed prosodic contours in natural language. When applied to question-answer pairs, the model straightforwardly predicts the well-established observation that words/phrases whose denotations are component meanings of every possible answer to the question under discussion are informationally redundant, and therefore do not bear prosodic prominence. In cases of contrastive focus on elements which are not so directly tied to a question under discussion, the model makes correct predictions if allowed to iterate the selection of prominent material throughout the syntactic structure of a sentence. In addition to providing a normative explanation for the emergence of observed focus systems, the game-theoretic model is extensible to problems related to focus association, and provides a probabilistic

[11] The following example, pointed out by an anonymous reviewer, highlights the primacy of linguistic salience over world-knowledge salience, but also introduces an interesting wrinkle.

(i) Q: Where does Sue live?
a. #She lives on MADISON avenue.
b. On Madison.

The sub-optimality of "MADISON avenue" is predicted by the primacy of linguistic salience in contexts where no prior avenues (as opposed to streets, etc.) have been mentioned. However, it is then surprising that *avenue* can be elided in the partial answer. The elision here is likely licensed by a peculiar property of street names and other sorts of proper nouns, and not by focus structure. These expressions can be shortened even when no part of the expression is in focus.

(ii) Q: Which of our friends lives on Madison avenue?
A: SUE lives on Madison.

This suggests that the elision in these cases is independent of focus.

approach to handling the placement of contrastive foci when multiple contrasting antecedents are present.

In sum, the current analysis provides a promising approach to thinking about information structure in terms of strategically optimal communication. This supplies an additional level of explanation above that of existing formal descriptions. On the other side of the coin, the current approach serves to bring the phenomena of information structure under the umbrella of the growing field of game-theoretic pragmatics (Parikh 2001, Benz et al. 2006, Parikh 2010, Clark 2011), whose goal is to explore meaning in context using a well-established *lingua franca* for modeling human behavior and conventions.

References

Artstein, Ron. 2004. Focus below the word level. *Natural Language Semantics* 12.

Bacharach, Michael. 2006. *Beyond individual choice: Teams and frames in game theory*. Princeton, NJ: Princeton University Press.

Bacharach, Michael & Michele Bernasconi. 1997. The variable frame theory of focal points: An experimental study. *Games and Economic Behavior* 19(1). 1–45.

Bardsley, Nicholas, Judith Mehta, Chris Starmer & Robert Sugden. 2010. Explaining focal points: Cognitive Hierarchy Theory versus team reasoning. *The Economic Journal* 120(543). 40–79.

Beaver, David & Brady Clark. 2008. *Sense and sensitivity: How focus affects meaning*. Oxford: Blackwell.

Benz, Anton. 2012. Errors in pragmatics. *Journal of Logic, Language and Information* 21(1). 97–116.

Benz, Anton, Gerhard Jaeger & Robert van Rooij (eds.). 2006. *Game theory and pragmatics*. New York: Palgrave MacMillan.

Bergen, Leon & Noah Goodman. 2014. The strategic use of noise in pragmatic reasoning. In P. Bello, M. Guarini, M. McShane & B. Scassellati (eds.), *Proceedings of the 36th annual cognitive science conference*, Austin, TX: Cognitive Science Society.

Bolinger, Dwight. 1972. Accent is predictable (if you're a mind-reader). *Language* 48(3). 633–644.

Büring, Daniel. 2003. On D-trees, Beans, and B-accents. *Linguistics and Philosophy* 26(5). 511–545.

Büring, Daniel. 2008. What's new (and what's given) in the theory of focus? In *Proceedings of the berkeley linguistics society meeting*, University of California.

Camerer, Colin F., Teck-Hua Ho & Juin Kuan Chong. 2004. A cognitive hierarchy model of games. *The Quarterly Journal of Economics* 119(3). 861–898.

Chafe, Wallace. 1974. Language and consciousness. *Language* 50(1). 111–133.

Chomsky, Noam. 1971. Deep structure, surface structure and semantic interpretation. In D. Steinberg & L. Jakobovits (eds.), *Semantics: An interdisciplinary reader in philosophy*, 183–216. Cambridge: Cambridge University Press.

Chomsky, Noam. 2011. Language and other cognitive systems. what is special about language? *Language Learning and Development* 7(4). 263–278.

Clark, Robin. 2011. *Meaningful games*. Cambridge, MA: MIT Press.

De Kuthy, Kordula & Detmar Meurers. 2012. Focus projection, between theory and evidence. *Empirical Approaches to Linguistic Theory–Studies in Meaning and Structure* 111. 207–240.

De Kuthy, Kordula & Britta Stolterfoht. 2014. Focus projection revisited: Pitch accent perception in German. In *Presented at Linguistic Evidence 2014, Tübingen, Germany*, .

Féry, Caroline & Vieri Samek-Lodovici. 2006. Focus projection and prosodic prominence in nested Foci. *Language* 82(1). 131–150.

Franke, Michael. 2009. *Signal to act: Game theory in pragmatics*: Universiteit van Amsterdam dissertation.

Franke, Michael. 2011. Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantic and Pragmatics* 4. 1–82.

Fudenberg, Dan & Jean Tirole. 1991. Perfect Bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory* 53(2). 236–260.

German, James, Janet Pierrehumbert & Stefan Kaufmann. 2006. Evidence for phonological constraints on nuclear accent placement. *Language* 151–168.

Halliday, Michael. 1967. Notes on transitivity and theme in English, part II. *Journal of Linguistics* 3. 199–244.

Hamblin, C.L. 1973. Questions in Montague english. *Foundations of Language* 10. 41–53.

Harsanyi, John C. 1968. Games of incomplete information played by 'Bayesian' players, part ii. *Management Science* 14(5). 320–334.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar.* Cambridge, MA: MIT Press.

Katz, Jonah & Elizabeth Selkirk. 2011. Contrastive focus vs. discourse-new. *Language* 87(4). 771–816.

Krifka, Manfred. 2001. For a structured meaning account of questions and answers. In C. Fery & W. Sternefeld (eds.), *Audiatur Vox Sapientia: A Festschrift for Arnum von Stechow*, 287–319. Berlin: Akademie Verlag.

Kuno, Susumo. 1982. The focus of the question and the focus of the answer. *Parassession on Nondeclaratives. Chicago Linguistics Society* 134–157.

Ladd, D. Robert. 1996. *Intonational phonology.* New York: Cambridge University Press.

Lai, Catherine. 2012. *Rises all the way up: The interpretation of prosody, discourse attitudes and dialogue structure*: University of Pennsylvania dissertation.

Lewis, David. 1969. *Convention: A philosophical study.* Cambridge: Cambridge University Press.

Nash, John. 1950. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences* 36(1). 48–49.

Nash, John. 1951. Non-cooperative games. *Annals of mathematics* 54(2). 286–295.

Parikh, Prashant. 2001. *The use of language.* Stanford, CA: CSLI Publications.

Parikh, Prashant. 2010. *Language and equilibrium.* Cambridge, MA: MIT Press.

Prince, Ellen. 1981. Toward a taxonomy of given-new information. In P. Cole (ed.), *Radical pragmatics*, 223–256. New York: Academic Press.

Roberts, Craige. 1996. Information Structure: Towards an integrated formal theory of pragmatics. In J. Yoon & A. Kathol (eds.), *Osuwpl volume 49: Papers in semantics*, The Ohio State University.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1. 75–116.

Schelling, Thomas. 1960. *The strategy of conflict.* Cambridge, MA: Harvard University Press.

Schmitz, Hans-Christian. 2008. *Accentuation and interpretation.* New York: Palgrave MacMillan.

Schwarzschild, Roger. 1999. Givenness, AvoidF, and other constraints on the placement of accent. *Natural Language Semantics* 7. 141–177.

Selkirk, Elizabeth. 2007. Contrastive focus, givenness and the unmarked status of "discourse-new". In C. Fery, G. Fanselow & M. Krifka (eds.), *Interdisciplinary studies on information structure 6: The notions of information structure*, 125–145.

Shannon, Claude E. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27(3). 379–423.

Stalnaker, Robert. 1979. Assertion. *Syntax and Semantics* 3.

Stevens, Jon Scott. 2013. *Information structure, grammar and strategy in discourse.*: University of Pennsylvania dissertation.

Stevens, Jon Scott. 2014. Against a unified analysis of givenness and focus. In R.E. Santana-LaBarge (ed.), *Proceedings of the 31st west coast conference on formal linguistics*, Somerville, MA: Cascadilla Press.

Swerts, Marc, Emiel Krahmer & Cinzia Avesani. 2002. Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics* 30(4). 629–654.

Truckenbrodt, Hubert. 2007. The syntax-phonology interface. In P. de Lacy (ed.), *The Cambridge handbook of phonology*, Cambridge: Cambridge University Press.

Vallduví, Enric. 1990. *The informational component*: University of Pennsylvania dissertation.

Wagner, Michael. 2012a. Contrastive topics decomposed. *Semantics and Pragmatics* 5(8). 1–54.

Wagner, Michael. 2012b. Focus and givenness: A unified approach. In I. Kučerová & A. Neeleman (eds.), *Information structure: Contrasts and positions*, Cambridge: Cambridge University Press.

Zimmermann, Malte & Edgar Onea. 2011. Focus marking and focus interpretation. *Lingua* 121(11). 1651–1670.