HAL
open science

The role of speaker beliefs in determining accent placement

James German, Eyal Sagi, Stefan Kaufmann, Brady Clark

To cite this version:

James German, Eyal Sagi, Stefan Kaufmann, Brady Clark. The role of speaker beliefs in determining accent placement. A. Benz; C. Ebert; G. Jäger; R. van Rooij. Language, games, and evolution, Springer-Verlag, pp.92-116, 2011, 978-3-642-18006-4. 10.1007/978-3-642-18006-4_5 . hal-01482602

HAL Id: hal-01482602

https://hal.science/hal-01482602v1

Submitted on 22 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Role of Speaker Beliefs in Determining Accent Placement

James German*, Eyal Sagi†, Stefan Kaufmann†‡, Brady Clark†

*Laboratoire Parole et Langage
†Northwestern University
‡University of Göttingen

April 5, 2010

1 Introduction

In English and other languages, the distribution of nuclear pitch accents[1] within a sentence usually reflects how the meaningful parts of the sentence relate to the context. Generally speaking, the nuclear pitch accent can only occur felicitously on *focused* parts of the sentence, corresponding to information that is not contextually retrievable or *given*.[2] In most contemporary theories, focus is formally represented by an abstract syntactic feature 'F'. Those parts of the sentence that are *given* tend to resist F-marking and thus nuclear accentuation.[3] In short, there is a more or less tight coupling between (i) the contextual *information status* of parts of the sentence; (ii) the *focus structure* of the sentence (represented by the distribution of syntactic F-marking); and (iii) the actual *accent placement* in the phonological form.

[1] This is typically defined as the last pitch accent in an intermediate phrase (Pierrehumbert, 1980; Beckman and Pierrehumbert, 1986).

[2] We follow Schwarzschild's suggestion to abandon the use of *new* as a collective technical term for the complement of *given* material, since doing so would intimate a homogeneity that is not there.

[3] There are cases in which other constraints overrule this tendency and force prosodic pominence on given material (see Schwarzschild, 1999, for examples and discussion).

1.1 Focus projection and the grammar

A commonly encountered view holds that the relation between information status, F-marking, and accent is governed by rules that may be complicated, but are nonetheless unequivocal, even deterministic, so that relative to a particular context exactly one placement of the accent is felicitous. One can easily adduce data for which this is indeed the case, such as the question-answer pairs in (1).[4] In each of B's responses, the part which corresponds to the '*wh*'-phrase of the question is the focus of the answer and the only natural location of the nuclear pitch accent. Consequently, the same syntactic string has to be pronounced differently in response to different questions: While each of the answers in (1a-c) is felicitous in the context of its question, it cannot be felicitously replaced with either of the others.

(1) a. A: Who did John praise?
B: John praised MARY.
b. A: Who praised Mary?
B: JOHN praised Mary.
c. A: What did John do to Mary?
B: John PRAISED Mary.

In general, though, the correspondence between accent placement and the questions an utterance can felicitously answer is not so tight. Selkirk (1996) noted that (2), in which the accent is located on the word '*bats*', can answer any of the questions in (2a-e).

(2) Mary bought a book about BATS.
a. What did Mary buy a book about?
b. What kind of book did Mary buy?
c. What did Mary buy?
d. What did Mary do?
e. What's been happening?

Depending on which question the sentence is used to answer, different constituents of it are in focus (those which correspond to the '*wh*'-phrases of the respective questions), and while all of these focused constituents contain the accented word '*bats*', they also contain additional, unaccented material in all cases except (2a).

[4] As usual in the literature, the location of the nuclear pitch accent is typographically indicated by capitals. This convention glosses over certain details of the intonation contour and is therefore not always appropriate. It is sufficient for our purposes, however.

Selkirk and others have argued that the relationship between accent and focus is mediated through an abstract syntactic feature F. This feature must originate from an accented word, but may percolate to other constituents, subject to certain syntactic rules of *focus projection* (Chomsky, 1972). In Selkirk's system, F-marking may spread (i) from an internal argument to its head, and (ii) from a head to the constituent it projects. Thus in (3) the word *'bats'* must be accented if it is to be F-marked at all, since for syntactic reasons the F-marking could not project to *'bats'* from any other location in the sentence. However, the F-marking *can* project from *'bats'* to each of the constituents listed as F-marked in (3b-e). The corresponding questions from (2) are given on the right.

(3)	a.	Mary bought a book about $[\text{BATS}]_{\text{F}}$.	(2a)
	b.	Mary bought a book $[\text{about BATS}]_{\text{F}}$.	(2b)
	c.	Mary bought $[\text{a book about BATS}]_{\text{F}}$.	(2c)
	d.	Mary $[\text{bought a book about BATS}]_{\text{F}}$.	(2d)
	e.	$[\text{Mary bought a book about BATS}]_{\text{F}}$.	(2e)

In most accounts of focus projection, the grammatical rules fully determine where the accent must fall in order to realize a given focus structure.[5] Once the latter is fixed, there is no room for variation, let alone speaker choice.

1.2 The role of speaker choice

Two steps lead from context to accent: Contextual information status maps to focus structure, which in turn guides and constrains the placement of nuclear pitch accents. Both of these steps are frequently treated as though they were governed by deterministic grammatical principles. This seems to be an oversimplification at both levels, however.

First, regarding the relationship between information status and focus structure, some authors explicitly assume that focus is part of speakers' communicative intentions, thus representing an active choice. For Roberts (1996), for example, the focus of a sentence indicates which question or issue the speaker takes to be the one currently under discussion; in this sense, focus is a means of keeping interlocutors' common ground and communicative goals in alignment. For Schwarzschild (1999), focus structure is determined by a particular type of anaphoric relationship between parts of an utterance

[5] Kadmon (2001) and Winkler (1997) contain recent major overviews of the literature on focus projection.

and the contents of the discourse context, which he calls *Givenness*. Here, speakers have some freedom in the choice of anaphoric relationships.

Second, under certain conditions, syntactic rules constraining accent placement (such as those presented in Selkirk, 1996) are violated. German et al. (2006) showed that speakers tend to avoid placing nuclear accents on prepositions, even in contexts in which those prepositions are new and the only alternative is to place an accent on given material. Thus when uttering (4B) in the context of (4A), speakers prefer to place the nuclear accent on the direct object *'game'* (5a) rather than the preposition *'in'* (5b), even though the preposition is the only new information in the clause.

(4) A. I noticed that Liz and Sally really like to play their game.
B. Unfortunately, Paul wrecked the tent that they play their game in.

Interestingly, though, the avoidance of accented prepositions was only partial in their study. While the overall preference was for patterns like (5a), speakers also produced the pattern in (5b), which accords with Selkirk's focus projection rules. The design of the experiment ensured that this variability could not be attributed to speaker or experimental error.

(5) a. ...that they play their GAME in.
b. ...tent that they play their game IN.

German et al. account for this finding by introducing an OT-style markedness constraint into the grammar which militates against forms with nuclear accents on prepositions. To explain the variability in outcomes, they follow Anttila (1997) and Boersma and Hayes (2001) in proposing that this markedness constraint interacts probabilistically with other constraints governing the distribution of focus in the general case.

This treatment may provide as good an account of the variation as one can expect from the constraints considered by German et al. (2006), but there are reasons to doubt that it actually *explains* what is going on. The observation is that speakers can use a form (the one with the accent on the direct object) with a focus structure with which it is not conventionally associated. Modifying the grammatical principles to accommodate this fact would seem to imply that the form in question can in some sense *mean* the same as the one that is conventionally associated with the focus structure in question. But intuitively, the deviating form is "pressed into service," so to speak, despite the fact that it does *not* mean the same.

In this paper, we propose instead to treat focus projection rules such

as those presented in Selkirk (1996) as just one factor among several that influence accent placement, and ask whether and under what circumstances it may be safe or even advantageous for speakers to violate those rules. Specifically, we conjecture that the observations of German et al. are not due to random variation after all, but rather to factors which were not represented in their model and, consequently, not controlled for in their experiment.[6]

In a nutshell, our proposal is that speakers' beliefs about hearers' expectations play a role in determining when to use certain accent patterns. On the one hand, speakers' tendency to avoid accenting prepositions is due to a *cost* associated with the effort involved in using such a form. On the other hand, in certain contexts the hearer can "guess" the information structure of a sentence independently of the accent pattern it carries. In cases like (4), the speaker's choice comes down to a tradeoff: If there is a substantial risk that the hearer would choose the wrong interpretation without the information carried by the accent pattern, then the speaker will pay the extra cost and accent the preposition. If the risk of miscommunication is low, however, the speaker will tend to avoid accenting the preposition.

We formalize this tradeoff in a signaling game and explore the predictions resulting in terms of either Nash equilibrium strategies or Pareto-Nash dominant strategies (Parikh, 2001, 2010). Importantly, which strategy dominates is predicted to depend on the prior probabilities of the various information-structural interpretations under consideration. In contrast to earlier treatments of information structure and accent placement, our model does not deal with the structure of the grammar directly, but with the extent to which speakers and hearers are bound by the grammar in negotiating their respective communicative goals and preferences. Our application differs from previous game-theoretic treatments (such as those in Parikh, 2010; Benz et al., 2006) in that it is not merely concerned with pragmatic enrichment or strengthening, but with a case in which winning strategies may step outside the form-meaning mappings licensed by the grammar.

In Section 2, we discuss in some more detail the main assumptions and intuitions underlying our proposal. Section 3 presents the formal version of our model as well as its application to the key problem that this paper addresses. Section 4 discusses the main results and implications, and sec-

[6]There are precedents for the view that pragmatic factors may override the rules of grammar, for instance in binding theory. Thus Chomsky (1981) argues that "... these contexts [e.g. contexts that license Principle C violations] do not constitute counterevidence to principle (C); rather they indicate that principle (C) may be overridden by some condition on discourse, not a very startling fact."

tion 5 concludes with some brief remarks about how the approach might be broadened and carried forward.

2 Contextual factors in accent placement

What are the factors that may be driving speakers' choices in cases like (4B)? Central to our proposal is the assumption that extant theories of focus and accent placement are essentially right (e.g., that the accent "belongs" on the preposition in (4B)), and that the role of information structure in determining accent placement is part of the knowledge speakers and hearers bring to their interactions.

In terms of its communicative function, the placement of accents is significant for a variety of reasons. For instance, it has been argued that by placing prosodic prominence on those parts of an utterance which introduce new information, speakers draw hearers' attention to those parts and facilitate their understanding (Schmitz, 2005). Aside from this facilitating role, accent placement is also an aide in synchronizing speakers' and hearers' respective beliefs about the common ground and the goals of the ongoing interaction. Seen this way, accent placement is a *grounding* device (Clark and Wilkes-Gibbs, 1986; Clark, 1996; see Thompson, 2009 for an overview and references).

Under what circumstances, then, would we expect the speaker to use a form that deviates from the grammatically "correct" one for the focus structure she has in mind? One such circumstance would be if that intended focus structure is highly expected independently of specific linguistic cues in the utterance itself – if, for example, the context already makes the intended focus structure highly salient or likely. In general, the more of a need the speaker feels to provide additional cues for her intended focus assignment, the stronger the incentive to use the form that the grammar licenses. Conversely, if the context provides strong cues for the intended focus assignment, accent placement loses its significance as a grounding device. Our model predicts that it is precisely in those situations that other factors – such as prosodic preferences – may outweigh focus structure in determining speakers' choices. On this view, it is not surprising that participants in German et al.'s study produced (5a), rather than (5b), in the context of (4A): Speakers feel free to use such "mismatched" forms whenever they can do so without risk of miscommunication.

Thus, one factor that we take to play an important role in determining speakers' choices is the hearer's uncertainty about the intended focus struc-

ture. In the formal model, this parameter is encoded as the hearer's subjective probability distribution over the various focus structures the speaker may have in mind. Following standard practice in the theory of signalling games, we assume that this probability distribution is common knowledge between both interlocutors.[7]

For the above example, this predicts that the speaker's license to produce the "mismatched" contour in (4B), with the accent placed on the direct object, in a situation in which she takes the direct object to be given, increases with the degree to which she believes its givenness to be already expected by the hearer. A case like (4B), where the expression *'their game'* directly repeats an expression from the previous utterance, leaves little uncertainty about the speaker's intended interpretation. Thus (4) is a very clear example of the situation we are describing: No miscommunication is likely to result from accenting the direct object.

This reasoning leads to predictions about the conditions under which speakers are allowed to produce accent patterns that do not match the intended focus structure. It does not yet explain, however, when and why speakers will actually do so. The fact that an incongruent accent placement is unlikely to result in miscommunication is not in itself a reason for preferring it over its congruent alternatives.

To address this question, we assume that in addition to the hearer's beliefs about the speaker's intended focus structure, three further factors play a role, all related to the *effort* involved in production and interpretation. In the model, they are represented as *costs* incurred by the interlocutors.

The first factor represents the speaker's effort in production. Here we follow German et al. as well as much of the literature on functional theories of grammar in invoking preferences against the production of particular forms (Croft, 1990; Eckman et al., 1986; Haiman, 1985, among others). Specifically, in our example, we assume that the congruent form with accentuation on the preposition *'in'* is dispreferred due to a general tendency against nuclear accentuation on certain function words (Ladd, 1980; Selkirk, 1995; German et al., 2006). Similarly to the optimality-theoretic approach, we

[7] This may seem to be an oversimplification, since the speaker does not really have access to the hearer's actual beliefs and may be mistaken about them. But recall that our goal is to model the factors which motivate the speaker's choice of a form, at the time she makes her utterance. Although the speaker takes the listener's perspective into consideration in making her decision, her choice can only be informed by what she takes to be the hearer's beliefs, not by the hearer's actual beliefs. Therefore, although one could devise a more complicated model which allows for the possibility that the speaker is wrong about the hearer's actual probability distribution over various focus structures, this extension would not contribute substantially to the analysis we are concerned with.

assume that this tendency is always operative – thus the cost is always incurred by speakers who produce the accent on the preposition – but may be outweighed by other forces.

The remaining two factors arise from a mismatch between the grammatically determined focus-to-accent mapping and the actual choices made in production and interpretation. On the one hand, the speaker incurs a cost whenever she chooses an accent pattern that is not the one specified by the grammar for the focus structure she has in mind. Essentially, the speaker prefers, all else being equal, to adhere to the grammar, and will not deviate spuriously. Similarly, the hearer incurs a cost whenever he chooses, as an interpretation, a focus structure that is not grammatically consistent with the accent pattern that the speaker has produced. Like the speaker, then, the hearer prefers not to deviate from the grammar, but may do so when other considerations apply.

We might view this extra cost, particularly as it concerns the hearer, as a *processing cost* incurred by extra inferences required to decide whether such a mismatch should be permitted in the given situation. A different, though perhaps not unrelated role of this cost would be in perpetuating the grammatical system throughout the population as well as diachronically. Without some formal reflex of such "grammaticality-bias" in the model, there may be no particular advantage to any one pairing between forms and meanings. Such a framework would furnish an account that we consider bizarre, namely that the grammatical role played by nuclear accent is itself subject to variation: Depending on the contextually given probabilities, accent sometimes marks non-givenness (as is standardly the case) and sometimes givenness. Intuitively, one of these uses is the norm from which the other deviates. The cost we stipulate is intended to account for this intuition.

Overall, interlocutors share the goal of successful communication, and both will assign a positive value to combinations of actions that lead to the hearer's choosing the information structure that the speaker had in mind. Speakers and hearers also prefer choices which adhere to the grammatically specified correspondence between accent patterns and information structure, and costs are incurred whenever this correspondence is contravened. Finally, speakers may incur an additional cost for using particular forms. For our examples, it is sufficient to assume that, all else being equal, utterances that include nuclear accents on prepositions incur a greater cost than ones with nuclear accents on full nouns.

Apart from the costs and benefits, our model assumes that interlocutors share certain beliefs regarding which focus structure a speaker is likely to convey for a given sentence in a given context. Formally, this is a function

from (explicit) contexts to functions from sentences to probability distributions over possible focus structures. In the cases we discuss, the context and textual content of the sentence are known. Thus, for practical purposes, this feature can be reduced to a single, mutually accessible probability distribution over those focus assignments that are sensible given the syntax and lexical content of the sentence being uttered.

All of this fundamentally assumes that the grammar provides a fixed mapping between information structure and accent placement. Yet we have not yet specified which version of such a grammar we are assuming. In fact, for the purposes of our analysis, a few minimal assumptions suffice. We take the core aspects of Schwarzschild (1999) as the foundation of our simplified theory. Specifically, we assume that (i) all non-F-marked nodes in the syntax are interpreted as given, and (ii) each given node introduces a presupposition that there is an antecedent in the context with which it co-refers.[8] Finally, we assume that nuclear accentuation introduces F-marking, which relieves the constituents in question of the givenness presupposition.[9]

In the German et al. study, the vast majority of productions broke down in two basic categories: one in which the only nuclear accent in the embedded clause falls on the direct object, as illustrated in (5a), and one in which it falls on the stranded preposition, as in (5b).

(5) a. ...that they play their GAME in.
b. ...that they play their game IN.

There were several prosodic variations in the material preceding the embedded clause (the most common being whether the head of the relative clause received an accent), but these differences did not substantially affect the predictions that Selkirk (1996) and Schwarzschild (1999) make for the embedded clause itself. Specifically, these theories predict that a pattern like (5a) can realize a number of focus assignments, including the ones in (6): In (6a) the direct object *'their game'* is treated as the only F-marked element, whereas in (6b) the F-marking projects to the entire verb phrase.

(6) ...they play their GAME in
a. ...they play $[\text{their game}]_{\text{F}}$ in
b. ...they $[[[\text{play}]_{\text{F}}\ [\text{their game}]_{\text{F}}]_{\text{F}}\ [\text{in}]_{\text{F}}]_{\text{F}}$

[8] Schwarzschild in fact allows for a more inclusive notion of inferability, formally modeled in terms of entailment under existential closure. In the minimal contexts we consider, this relation does not add anaphoric possibilities beyond those available by coreference.

[9] This does not mean that the element *must not* have a contextually salient antecedent, but merely that it carries no presupposition to that effect.

While (5a) is thus consistent with a number of focus structures, all of them have in common that the direct object *'their game'* is treated as F-marked. Since that is the feature we are most interested in, we ignore the differences and collapse all of these cases into one. In contrast, for a pattern like (5b) the theories predict that the preposition *'in'* must be interpreted as the only F-marked constituent in its clause; in particular, the pattern is predicted to be inconsistent with F-marking on the direct object.

Thus we draw the relevant distinction in terms of F-marking on the direct object *vs.* the preposition and adopt the notation in (7). It should be reiterated, however, that this is merely a shorthand notation and that the alignment between accent placement and F-marking is more complicated.

(7) a. ...they play [their game]$_{\mathrm{F}}$ in (5a)
b. ...they play their game [in]$_{\mathrm{F}}$ (5b)

With these preliminaries in place, we turn to the specification of our formal game-theoretic model.

3 Formal model

For our purposes, a *language* consists of two non-empty sets F (of *forms*) and M (of *meanings*). Since we are interested in a language which comes with a conventional interpretation, the formal model should also include a mapping of some kind between F and M, such as a relation R in $F \times M$, which constrains the interpretations available for each form. But since our main point is that speakers can and do "step outside" the bounds imposed by the conventional interpretation, the conventional interpretation should be capable of interacting with and being overruled by other forces. To this end, we assume that the interpretation is given as part of the payoff structure – specifically, in terms of *costs* of production and interpretation.

Our example involves the speaker's choice between the two forms in (5a) and (5b). Here we label them f_{NP} and f_{np}, indicating the placement of the nuclear pitch accent on the noun phrase or on the preposition, respectively. The two meanings we are concerned with are m_{N} and m_{G}, corresponding to the focus-structural status of the noun phrase (see (7a) and (7b) above). Thus under the familiar grammatical constraints on accent placement, the pairings $\langle m_{\mathrm{N}}, f_{\mathrm{NP}}\rangle$ and $\langle m_{\mathrm{G}}, f_{\mathrm{np}}\rangle$ are congruent, whereas $\langle m_{\mathrm{N}}, f_{\mathrm{np}}\rangle$ and $\langle m_{\mathrm{G}}, f_{\mathrm{NP}}\rangle$ are "mismatches."

In our game-theoretical model, a speaker strategy is a function σ mapping meanings to forms, and a hearer strategy is a function τ mapping forms

to meanings. On each occasion of use, the speaker and the hearer choose *strategy profile.*

Definition 1 (Strategies and strategy profiles)**.** *Let a language* $\mathcal{L} = \langle F, M \rangle$ *be given. The set of* speaker strategies *for* $\mathcal{L}$*,* $\mathcal{S}_s$*, is the set of functions* $\sigma : M \mapsto F$*. The set of* hearer strategies *for* $\mathcal{L}$*,* $\mathcal{S}_h$*, is the set of functions* $\tau : F \mapsto M$*. The set of* strategy profiles *is* $\mathcal{S} = \mathcal{S}_s \times \mathcal{S}_h$*.*

The speaker utters the form which her strategy assigns to the meaning she wants to convey, and the hearer uses his strategy to map the form he receives to a meaning. We assume that the form the speaker utters and the one the hearer perceives are identical, thus there is no noise. Since σ and τ are functions, once they are fixed, the outcome of the exchange is determined by the speaker's intended meaning.

Definition 2 (Costs and benefits)**.** *Let* $\mathcal{L} = \langle F, M \rangle$*. A function* $C^s_{\mathrm{P}} : F \mapsto \mathbb{R}$ *assigns to each form in* F *a cost incurred by the speaker for uttering it. Two functions* $C^s_{\mathrm{M}}, C^h_{\mathrm{M}} : (F \times M) \mapsto \mathbb{R}$ *assign to each form-meaning pair* $\langle f, m \rangle$ *a cost incurred by the speaker and the hearer, respectively, for producing and interpreting* f *as conveying* m*. The benefit of successful communication is given by a function* $B : (M \times M) \mapsto \mathbb{R}$*, such that for each* $m, m' \in M, B(m, m') > 0$ *if* $m = m'$*, 0 otherwise.*

In our example, the costs are represented by variables as follows. The production costs for the forms in question are C^s_{np} for placing the accent on the preposition, and C^s_{NP} for placing it on the noun phrase. These production costs arise due to prosodic constraints governing the respective forms and are independent of the information structure. Furthermore, $C^s_{\text{✗}}$ and $C^h_{\text{✗}}$ represent the speaker's and hearer's respective costs of producing and processing a "mismatched" interpretation of a nuclear pitch accent (i.e., mapping the accented constituent to given information). In contrast, $C^s_{\checkmark}$ and $C^h_{\checkmark}$ are the respective costs of producing and processing the "canonical" pairings which map the accented constituent to new information. Based on the above discussion, we take it that generally $C^s_{\mathrm{np}} > C^s_{\mathrm{NP}}$, $C^s_{\text{✗}} > C^s_{\checkmark}$, and $C^h_{\text{✗}} > C^h_{\checkmark}$. As we will see, the relative magnitude between these pairs of costs is more important than their absolute values.

Successful communication is rewarded by a benefit which we stipulate is positive if the meaning the hearer extracts is the same as the one the speaker intended to convey (thus communication is successful), and zero otherwise. For all possible outcomes, the benefit is the same for both interlocutors. Thus the game is one of *coordination.* This choice rules out many real-life

situations, such as ones in which the speaker has an interest in misleading the hearer. The predictions of the model would change considerably in such cases, but we exclude them here because such situations lie beyond the purview of this paper.

For each linguistic encounter, the benefits and costs associated with the chosen strategy profile jointly determine its utility:

Definition 3 (Utility)**.** *Given C^s_{P}, C^s_{M}, C^h_{M} and B, a* utility *function $U : M \times F \times M \mapsto \mathbb{R}$ for $\mathcal{L}$ is defined as follows: For all $m, m' \in M$ and $f \in F$,*

$$U(m, f, m') = B(m, m') - C^s_{\mathrm{P}}(f) - C^s_{\mathrm{M}}(f, m) - C^h_{\mathrm{M}}(f, m')$$

Now, neither of the interlocutors knows the other's choice of strategy, and the hearer only has probabilistic information about the speaker's intended meaning. Therefore the outcome is not predictable with certainty. However, since the hearer's beliefs about the speaker's intentions are common knowledge, both participants are able to calculate the *expected utility* of each strategy pair $\langle \sigma, \tau \rangle$ – the weighted sum of the utilities for each of the meanings the speaker may intend, where the weights are the hearer's subjective probabilities of those meanings.

Definition 4 (Expected utility)**.** *Let $\mathcal{L} = \langle F, M \rangle$ be a language, U a utility function for $\mathcal{L}$ and $P : M \mapsto [0, 1]$ a probability distribution over the meanings in $\mathcal{L}$ such that for each $m \in M$, $P(m)$ is the hearer's prior probability that the speaker intends to convey m. The* expected utility *for $\mathcal{L}$ given U and P is a function $EU : \mathcal{S} \mapsto \mathbb{R}$ defined as follows, for all $\langle \sigma, \tau \rangle \in \mathcal{S}$:*

$$EU(\sigma, \tau) = \sum_{m \in M} P(m) \times U(m, \sigma(m), \tau(\sigma(m)))$$

With the definitions so far, we have secured all the ingredients for a *game* in the formal sense.

Definition 5 (Game)**.** *Given a language $\mathcal{L}$, a utility function U as defined above, and a probability distribution P over meanings in $\mathcal{L}$, a (two-player)* game *for $\mathcal{L}$ is a triple $\Game = \langle \{s, h\}, \mathcal{S}, EU \rangle$, where s, h are speaker and hearer, respectively; $\mathcal{S}$ is the set of strategy profiles for $\mathcal{L}$; and EU is the expected utility function for $\mathcal{L}$ given U and P.*

The most fundamental and commonly used notion in making this prediction is that of a *Nash Equilibrium*. In a game of coordination like ours, there is always at least one Nash Equilibrium; in general, there may be more than one.

Definition 6 (Nash Equilibria)**.** *The set of* Nash equilibria *in a game* $\Game$ *is the set*

$$NE(\Game) = \{\langle\sigma,\tau\rangle | \forall\sigma'[EU(\sigma',\tau) \leq EU(\sigma,\tau)] \\ \wedge\forall\tau'[EU(\sigma,\tau') \leq EU(\sigma,\tau)]\}$$

The Nash equilibrium has been employed in linguistic analyses by Lewis (1969), Dekker and van Rooij (2000), and others. However, it has some limitations which have prompted some authors to look for refinements and alternatives. Parikh's (2001, Section 4.4; 2010, Section 3.3.5) proposal is to filter the Nash equilibria in a given game by the criterion of *Pareto dominance* in order to eliminate "local minima" and retain only those that are closer to our intuitive notion of "best choice." Overall, the question of appropriate solution concepts for various kinds of games is still open (cf. van Rooij, 2004:506; Parikh, 2006). Here we adopt Parikh's strategy of Pareto-dominance as the criterion for the normative model. More specifically, we adopt the notion of *weak* Pareto-dominance, which ensures that if there is at least one Nash Equilibrium in the game, then there is a (not necessarily unique) Pareto-dominant one. Therefore, since our games are guaranteed to have Nash Equilibria, at least one of them has to be Pareto-dominant.

Definition 7 (Pareto-Nash Equilibria)**.** *The set of* Pareto-Nash Equilibria *in a game* $\Game$ *is the set*

$$PNE(\Game) = \{\langle\sigma,\tau\rangle | \forall\sigma',\tau'[EU(\sigma,\tau) \leq EU(\sigma',\tau') \\ \rightarrow EU(\sigma,\tau) = EU(\sigma',\tau')]\}$$

With these formal notions in place, let us now examine our example more closely. Recall that the set of meanings is $\{m_{\mathrm{N}}, m_{\mathrm{G}}\}$ (where the noun phrase is new and given, respectively) and the set of forms is $\{f_{\mathrm{NP}}, f_{\mathrm{np}}\}$ (where the accent is placed on the noun phrase or on the preposition). Table 1 lists all possible speaker and hearer strategies together with the associated costs for speaker and hearer for each possible move.

4 Results

In this section we present an analysis of the model in terms of the dominance relationships between the expected utilities of strategy sets under various conditions.[10] We have limited our detailed analysis to just the first

[10]The relationships between the various strategy sets take the form of conditions on dominance based on the variables that the model includes. In certain cases, these con-

Table 1: Speaker and hearer strategies with costs incurred by each move

Speaker			*Hearer*	
strategies	*costs*		*strategies*	*costs*
$\sigma_1 : \begin{bmatrix} m_{\text{N}} \mapsto f_{\text{NP}} \\ m_{\text{G}} \mapsto f_{\text{np}} \end{bmatrix}$	C^s_{NP}	$C^s_{\checkmark}$	$\tau_1 : \begin{bmatrix} f_{\text{NP}} \mapsto m_{\text{N}} \\ f_{\text{np}} \mapsto m_{\text{G}} \end{bmatrix}$	$C^h_{\checkmark}$
	C^s_{np}	$C^s_{\checkmark}$		$C^h_{\checkmark}$
$\sigma_2 : \begin{bmatrix} m_{\text{N}} \mapsto f_{\text{NP}} \\ m_{\text{G}} \mapsto f_{\text{NP}} \end{bmatrix}$	C^s_{NP}	$C^s_{\checkmark}$	$\tau_2 : \begin{bmatrix} f_{\text{NP}} \mapsto m_{\text{G}} \\ f_{\text{np}} \mapsto m_{\text{G}} \end{bmatrix}$	$C^h_{\times}$
	C^s_{NP}	$C^s_{\times}$		$C^h_{\checkmark}$
$\sigma_3 : \begin{bmatrix} m_{\text{N}} \mapsto f_{\text{np}} \\ m_{\text{G}} \mapsto f_{\text{NP}} \end{bmatrix}$	C^s_{np}	$C^s_{\times}$	$\tau_3 : \begin{bmatrix} f_{\text{NP}} \mapsto m_{\text{G}} \\ f_{\text{np}} \mapsto m_{\text{N}} \end{bmatrix}$	$C^h_{\times}$
	C^s_{NP}	$C^s_{\times}$		$C^h_{\times}$
$\sigma_4 : \begin{bmatrix} m_{\text{N}} \mapsto f_{\text{np}} \\ m_{\text{G}} \mapsto f_{\text{np}} \end{bmatrix}$	C^s_{np}	$C^s_{\times}$	$\tau_4 : \begin{bmatrix} f_{\text{NP}} \mapsto m_{\text{N}} \\ f_{\text{np}} \mapsto m_{\text{N}} \end{bmatrix}$	$C^h_{\checkmark}$
	C^s_{np}	$C^s_{\checkmark}$		$C^h_{\times}$

two rows and columns – that is, to the strategy profiles involving σ_1, σ_2, τ_1 and τ_2. There are several reasons for this. First of all, we feel that the relationships between these strategy sets most clearly illustrate the intuitions behind the phenomenon we are modeling. $\langle \sigma_1, \tau_1 \rangle$, for example, represents the "canonical" situation in which the speaker and hearer fully observe the rules of the grammar, thereby maximizing the benefit from successful communication and minimizing the costs from grammatical mismatches, while $\langle \sigma_2, \tau_2 \rangle$ represents what is in many ways the most interesting deviation from this pattern: The speaker avoids the extra cost associated with accenting the preposition even when the NP is given, and the hearer interprets all forms as having a given NP.

Strategy sets involving σ_3, σ_4, τ_3 and τ_4 deviate in other, sometimes interesting ways. It should be noted, in fact, that under certain conditions, the set of Nash equilibria and even Pareto-dominant strategies is not limited to the first four strategy sets. In the discussion that follows our analysis

ditions are mathematically non-trivial, and may seem somewhat abstract in comparison to the concrete communicative processes that we are trying to model. It should be noted that we do not mean to imply that the variables in our model, to the extent that they have a psychological reality, take on a precise numerical value that one could measure with any precision. Nevertheless, the mathematical inequalities do serve to elucidate certain broad tendencies that are likely to hold if the factors we consider have any psychological reality at all, and these are discussed where appropriate.

(Section 4.2), we mention such cases and discuss their implications for our model. For reasons of space, however, we leave these out of the analysis itself, and we leave it to the reader to carry out the associated mathematical proofs.

4.1 Costs and cost differentials

First of all, in comparing strategy profiles we can dispense with using costs directly (e.g., C^s_{np} and C^s_{NP}) and operate with just the differences between them. The resulting rankings are the same because for any given strategy profile, the expected utility is just the weighted sum of the utility terms, where for each pair of related costs (e.g., C^s_{np} and C^s_{NP}), if one is incurred with probability x, then the other is incurred with probability $(1-x)$. Thus for instance, the total form cost incurred for any strategy set is described by the term in (8a), which is equivalent to (8b).

(8) a. $M \times C^s_{\text{np}} + (1 - M) \times C^s_{\text{NP}}$
b. $C^s_{\text{NP}} + M(C^s_{\text{np}} - C^s_{\text{NP}})$

Since the lower cost term (here, C^s_{NP}) is constant across strategy sets, it can be ignored, for it will always be subtracted out of any comparison or inequality between two strategy sets. In other words, pairs of terms like (8a) will henceforth be replaced by terms like (9), where D^s_{P} is the difference between the two costs.

(9) $M \times D^s_{\text{P}}$

With this in mind, we can write the payoff matrix as in Table 2. (For readability, the matrix is spread over two rows.)

Fact 1. *$\langle \sigma_1, \tau_1 \rangle$ dominates $\langle \sigma_1, \tau_2 \rangle$ whenever $P_{\text{G}} < 1$.*[11]

[11] *Proof.* Notice first that (i) and (iv) are equivalent: (ii) is obtained by substitution from the payoff matrix, the rest follows by simple algebra (recall that $B - P_{\text{G}}B = P_{\text{N}}B$).

(i) $EU(\langle \sigma_1, \tau_1 \rangle) > EU(\langle \sigma_1, \tau_2 \rangle)$

(ii) $B - P_{\text{G}}D^s_{\text{P}} > P_{\text{G}}B - P_{\text{G}}D^s_{\text{P}} - P_{\text{N}}D^h_{\text{M}}$

(iii) $P_{\text{N}}B + P_{\text{N}}D^h_{\text{M}} > 0$

(iv) $P_{\text{N}}(B + D^h_{\text{M}}) > 0$

Clearly (iv) is true if and only if both factors on the left-hand side are true; i.e., iff $P_{\text{N}} > 0$ (equivalently, $P_{\text{G}} < 1$) and $B + D^h_{\text{M}} > 0$. The latter is assumed. □
We omit the proofs of subsequent results; they are obtained in a similar fashion.

Table 2: Payoff matrix

	τ_1	τ_2
σ_1	$B - P_{\mathrm{G}} D^s_{\mathrm{P}}$	$P_{\mathrm{G}} B - P_{\mathrm{G}} D^s_{\mathrm{P}} - P_{\mathrm{N}} D^h_{\mathrm{M}}$
σ_2	$P_{\mathrm{N}} B - P_{\mathrm{G}} D^s_{\mathrm{M}}$	$P_{\mathrm{G}} B - P_{\mathrm{G}} D^s_{\mathrm{M}} - D^h_{\mathrm{M}}$
σ_3	$-P_{\mathrm{N}} D^s_{\mathrm{P}} - D^s_{\mathrm{M}}$	$P_{\mathrm{G}} B - P_{\mathrm{N}} D^s_{\mathrm{P}} - D^s_{\mathrm{M}} - P_{\mathrm{G}} D^h_{\mathrm{M}}$
σ_4	$P_{\mathrm{G}} B - D^s_{\mathrm{P}} - P_{\mathrm{N}} D^s_{\mathrm{M}}$	$P_{\mathrm{G}} B - D^s_{\mathrm{P}} - P_{\mathrm{N}} D^s_{\mathrm{M}}$
	τ_3	τ_4
σ_1	$-P_{\mathrm{G}} D^s_{\mathrm{P}} - D^h_{\mathrm{M}}$	$P_{\mathrm{N}} B - P_{\mathrm{G}} D^s_{\mathrm{P}} - P_{\mathrm{G}} D^h_{\mathrm{M}}$
σ_2	$P_{\mathrm{G}} B - P_{\mathrm{G}} D^s_{\mathrm{M}} - D^h_{\mathrm{M}}$	$P_{\mathrm{N}} B - P_{\mathrm{G}} D^s_{\mathrm{M}}$
σ_3	$B - P_{\mathrm{N}} D^s_{\mathrm{P}} - D^s_{\mathrm{M}} - D^h_{\mathrm{M}}$	$P_{\mathrm{N}} B - P_{\mathrm{N}} D^s_{\mathrm{P}} - D^s_{\mathrm{M}} - P_{\mathrm{N}} D^h_{\mathrm{M}}$
σ_4	$P_{\mathrm{N}} B - D^s_{\mathrm{P}} - P_{\mathrm{N}} D^s_{\mathrm{M}} - D^h_{\mathrm{M}}$	$P_{\mathrm{N}} B - D^s_{\mathrm{P}} - P_{\mathrm{N}} D^s_{\mathrm{M}} - D^h_{\mathrm{M}}$

In descriptive terms, this just means that whenever the speaker is using a strategy that is sensitive to her intended meaning and accords with the rules of focus projection, then it is always better if the hearer uses a strategy that is sensitive to form and also accords with those rules.

Fact 2. *$\langle \sigma_1, \tau_1 \rangle$ dominates $\langle \sigma_2, \tau_1 \rangle$ whenever $P_{\mathrm{G}} > 0$ and $D^s_{\mathrm{P}} < B + D^s_{\mathrm{M}}$.*

Since we are assuming that D^s_{M} is positive, under the further assumption that $D^s_{\mathrm{P}} < B$ this condition will always be met.

In fact, there is good reason to assume that $D^s_{\mathrm{P}} < B$ as a general fact. If the cost of accenting a preposition were greater than the benefit that could be gained from successful communication, then it would always be better to remain silent than to produce such a form. This is not what is observed, however. Speakers in the German et al. study, especially, did accent prepositions, so the cost of doing so cannot be higher than the maximum benefit that can be attained in this way.

Intuitively, this result suggests that as long as that the hearer is using a strategy that is sensitive to form and conforms to the rules of focus projection, it is better for the speaker to mark her intentions in a way that also conforms to the grammar. Avoiding the cost of accenting the preposition will never sufficiently offset the risk of unsuccessful communication in such a case.

Fact 3. *$\langle \sigma_2, \tau_2 \rangle$ dominates $\langle \sigma_2, \tau_1 \rangle$ whenever $P_{\mathrm{G}} > (D^h_{\mathrm{M}}/2B) + 0.5$.*

As long as $D^h_{\text{M}} < B$, the right side of the inequality ranges between 0.5 and 1.[12] This means that P_{G} must be at least as large as 0.5 in order for the condition to hold. In addition, for a fixed B, the minimum condition on P_{G} increases linearly as a function of D^h_{M}.

Since the speaker always accents the NP in σ_2, a hearer using τ_2 will always incur a mismatch cost, regardless of the speaker's actual intention. This result then suggests that the probability of the NP being given (P_{G}) has to be high enough so that the increased chance of successful communication is sufficient to offset the hearer's cost of always deviating from the grammar. If that probability is too low, or D^h_{M} is too high, then it would be better for the hearer to interpret accented NPs as new, and accept miscommunication in all cases where the speaker intends the NP to be given.

Fact 4. *$\langle \sigma_2, \tau_2 \rangle$ dominates $\langle \sigma_1, \tau_2 \rangle$ whenever $D^s_{\text{P}} > D^s_{\text{M}} + D^h_{\text{M}}$ and $P_{\text{G}} > 0$.*

If the hearer is using an insensitive interpretive strategy, then it is preferable that the speaker use a uniform marking strategy whenever the cost of accenting the preposition is higher than either the speaker or hearer mismatch costs.

This relationship is less intuitive than the others, since it raises the question of why the speaker would bother to mark the focus structure with accent placement if the hearer is not attending to form. However, it makes more sense from the standpoint of cases where the NP is given. In those cases, using σ_2 always avoids the cost of accenting the preposition. However, since σ_2 and τ_2 each incurs a mismatch cost when the NP is given, the cost avoided (D^s_{P}) has to offset the costs incurred ($D^s_{\text{M}} + D^h_{\text{M}}$).

Fact 5. *$\langle \sigma_1, \tau_1 \rangle$ is a Nash equilibrium.*

This follows straightforwardly from Facts 1 and 2 and the definition of a Nash equilibrium. When $P_{\text{G}} = 0$ or $P_{\text{G}} = 1$, then this will be weakly true, since $EU(\langle \sigma_1, \tau_1 \rangle)$ will be equal to $EU(\langle \sigma_2, \tau_1 \rangle)$ and $EU(\langle \sigma_1, \tau_2 \rangle)$ respectively in those cases.

[12] See the discussion above regarding D^s_{P} and B in the preceding paragraphs. A similar argument applies here. If D^h_{M} were greater than B, then hearers would never deviate from the grammar of focus projection by interpreting an accented NP as given. They would adhere rigidly to the grammar even at a very high risk of miscommunication, in spite of any and all contextual evidence in favor of an ungrammatical interpretation. While the production data does not corroborate this assumption in the same way as for D^s_{P}, on an intuitive level, this is precisely what we are assuming licenses a speaker to contravene the rules of focus projection in the German et al. data.

Fact 6. *For* $P_{\mathrm{G}} > 0$, $\langle \sigma_2, \tau_2 \rangle$ *is a Nash equilibrium whenever* $P_{\mathrm{G}} > (D^h_{\mathrm{M}}/2B) + 0.5$ *and* $D^s_{\mathrm{P}} > D^s_{\mathrm{M}} + D^h_{\mathrm{M}}$.

This is just the conjunction of the conditions in Results 3 and 4. Note that when this stronger condition is met, then $\langle \sigma_1, \tau_1 \rangle$ and $\langle \sigma_2, \tau_2 \rangle$ are both Nash equilibria. When it is not met, then $\langle \sigma_1, \tau_1 \rangle$ is the only Nash equilibrium, since, except when $P_{\mathrm{G}} = 1$ or $P_{\mathrm{G}} = 0$, $\langle \sigma_2, \tau_2 \rangle$'s competitors, $\langle \sigma_1, \tau_2 \rangle$ and $\langle \sigma_2, \tau_1 \rangle$, are always dominated by $\langle \sigma_1, \tau_1 \rangle$.

Fact 7. $\langle \sigma_2, \tau_2 \rangle$ *strictly dominates* $\langle \sigma_1, \tau_1 \rangle$ *whenever* $P_{\mathrm{G}} > (B + D^h_{\mathrm{M}})/(B + D^s_{\mathrm{P}} - D^s_{\mathrm{M}})$.

Notice that since $P_{\mathrm{G}} \leq 1$, this can hold only if $D^h_{\mathrm{M}} \leq D^s_{\mathrm{P}} - D^s_{\mathrm{M}}$. This is equivalent to $D^s_{\mathrm{P}} \geq D^s_{\mathrm{M}} + D^h_{\mathrm{M}}$, which we know from Result 4 is a prerequisite for $\langle \sigma_2, \tau_2 \rangle$'s being equilibrium strategy in the first place. In addition, since we are assuming that $B > D^h_{\mathrm{M}}$ and $B > D^s_{\mathrm{P}}$, this condition implies that P_{G} must be at least greater than 0.5. Beyond that, the condition on P_{G} varies with (i) the difference between D^h_{M} and the term $(D^s_{\mathrm{P}} - D^s_{\mathrm{M}})$, and (ii) the magnitude of these two terms as a proportion of B. When D^h_{M} and $(D^s_{\mathrm{P}} - D^s_{\mathrm{M}})$ are relatively small as a proportion of B, then their difference will have little effect on the minimum condition for P_{G}, and that condition will be close to 1.0. By contrast, when those terms are large as a proportion of B, then their difference will have a large effect on the condition. When the difference is very small, then P_{G} must be close to 1.0, but as the difference increases, $\langle \sigma_2, \tau_2 \rangle$ may dominate $\langle \sigma_1, \tau_1 \rangle$ at smaller values of P_{G}.

This relationship is intuitively plausible, first of all, from the standpoint of the relative size of the factors as a proportion of the benefit for successful communication. When the difference between the cost of accenting a preposition (D^s_{P}) and the speaker mismatch cost (D^s_{M}) is very small, the switch from a strategy that avoids accenting prepositions (σ_2) to one that avoids mismatch costs in exactly the same cases (σ_1) is virtually an even trade, and there is little motivation to do so except when the probability is very high that the NP is given. This relationship also makes sense from the standpoint of speaker costs versus hearer costs. If the differential just described (i.e., $D^s_{\mathrm{P}} - D^s_{\mathrm{M}}$) is not much bigger than the cost a hearer incurs for a mismatch, then there is little motivation to use a strategy that incurs such a cost, except when it is very likely that the NP is given. By contrast, when the cost of accenting a preposition is very high as a proportion of B, the cost of both speaker and hearer mismatches are very low as a proportion of B, then it is desirable to use a pooling strategy that avoids production costs whenever it is even moderately more likely that the NP is given than not.

4.2 Discussion

Besides the equilibria described in Results 5, 6 and 7, there are several other interesting cases that lie outside of the four strategy profiles discussed above. To begin with, $\langle \sigma_2, \tau_3 \rangle$ is equivalent to $\langle \sigma_2, \tau_2 \rangle$, and therefore dominates its row in exactly the same set of cases as $\langle \sigma_2, \tau_2 \rangle$. This is intuitively clear when one considers that σ_2 only generates f_{NP}, which τ_2 and τ_3 both treat in the same way. The conditions under which $\langle \sigma_2, \tau_3 \rangle$ dominates in its column, however, are rather specific and unintuitive, and it is not of much use, therefore, to discuss the conditions under which $\langle \sigma_2, \tau_3 \rangle$ forms a Nash equilibrium. To the extent that it does, however, it cannot possibly be Pareto-dominant in any cases that $\langle \sigma_2, \tau_2 \rangle$ cannot also be. Moreover, when $\langle \sigma_2, \tau_2 \rangle$ and $\langle \sigma_2, \tau_3 \rangle$ are both (weakly) Pareto-dominant, it is not possible to distinguish between them behaviorally. In fact, one might speculate that speakers and hearers do not care which hearer strategy is being employed in such cases, since half of the strategy cannot possibly be relevant for the outcome.

It is also noteworthy that $\langle \sigma_2, \tau_4 \rangle$ is not only equivalent to $\langle \sigma_2, \tau_1 \rangle$, but also dominates its column under relatively weak sets of conditions. When the conditions described in Result 3 are not met, such that $\langle \sigma_2, \tau_1 \rangle$ dominates $\langle \sigma_2, \tau_2 \rangle$ (and $\langle \sigma_2, \tau_3 \rangle$), $\langle \sigma_2, \tau_4 \rangle$ may be a weak Nash equilibrium. Note, however, that according to Result 1, $\langle \sigma_1, \tau_2 \rangle$ is always dominated by $\langle \sigma_1, \tau_1 \rangle$ and can never itself be a Nash equilibrium. This implies even when $\langle \sigma_2, \tau_4 \rangle$ is a weak Nash equilibrium, it will be Pareto-dominated by $\langle \sigma_1, \tau_1 \rangle$.

Finally, $\langle \sigma_3, \tau_3 \rangle$ represents a surprisingly strong strategy set in our model under a range of conditions. It is a general result (whose proof we leave to the reader) that it dominates its own row whenever $0 < P_{\mathrm{G}} < 1$. It also dominates $\langle \sigma_1, \tau_3 \rangle$, and $\langle \sigma_4, \tau_3 \rangle$ whenever $B > D^s_{\mathrm{M}}$, which we are assuming anyway. Finally, it dominates $\langle \sigma_2, \tau_3 \rangle$, and therefore represents a Nash equilibrium, whenever $B < D^s_{\mathrm{M}} + D^s_{\mathrm{P}}$. Moreover, since $\langle \sigma_2, \tau_3 \rangle$ is equivalent to $\langle \sigma_2, \tau_2 \rangle$, $\langle \sigma_3, \tau_3 \rangle$ also Pareto-dominates $\langle \sigma_2, \tau_2 \rangle$ in those cases. Note however, that $EU(\langle \sigma_3, \tau_3 \rangle) > EU(\langle \sigma_1, \tau_1 \rangle)$ whenever $P_{\mathrm{G}} > (D^s_{\mathrm{M}} + D^h_{\mathrm{M}})/2D^s_{\mathrm{P}} + 0.5$. In other words, like $\langle \sigma_2, \tau_2 \rangle$, P_{G} must be somewhat higher than 0.5 before $\langle \sigma_3, \tau_3 \rangle$ even competes with $\langle \sigma_1, \tau_1 \rangle$ for Pareto-dominance.

This result suggests that using a strategy set like $\langle \sigma_2, \tau_2 \rangle$, that is insensitive to both intentions and forms, may not be the only rational alternative for avoiding costly forms. Under certain conditions, it may be better to use a strategy set that actually reverses the mapping associated with the rules of focus projection. This implies that both speaker and hearer mismatch costs are incurred for every possible outcome. This extra cost is offset, how-

ever, by the fact that communication is always successful. Furthermore, although $\langle \sigma_3, \tau_3 \rangle$ does not completely avoid extra form-based costs the way that $\langle \sigma_2, \tau_2 \rangle$ does, those costs play less of a role when the NP is very likely to be given, since the more costly form is the less likely to be used in such a case.

It is not possible to know from the German et al. results whether speakers who accented given NPs were using $\langle \sigma_2, \tau_2 \rangle / \langle \sigma_2, \tau_3 \rangle$ or $\langle \sigma_3, \tau_3 \rangle$. As already mentioned, $\langle \sigma_2, \tau_2 \rangle$ and $\langle \sigma_2, \tau_3 \rangle$ are indistinguishable from a behavioral standpoint, so there is no data that could ever distinguish between them. $\langle \sigma_3, \tau_3 \rangle$, on the other hand, predicts a distinct set of behavioral outcomes, which could, in principle, distinguish it from $\langle \sigma_2, \tau_2 \rangle$ and $\langle \sigma_2, \tau_3 \rangle$. We leave this to future research.

In many ways, however, there is something very counterintuitive about $\langle \sigma_3, \tau_3 \rangle$, even if our model suggests that is sometimes the most rational outcome. Does it make sense that speakers and hearers would, or ever do, temporarily negotiate a set of strategies that literally flips the grammar on its head? Temporarily resorting to a pooling strategy, on the other hand, is easier to imagine, and more closely resembles various other human behaviors (such as laziness) that have a stronger presence in popular discourse and folk psychology. Perhaps it is not unreasonable, then, to suppose that there are other biases involved that our model does not represent. In other words, while speakers and hearers may temporarily accept a slight deviation from the rules of the grammar when there are obviously costly forms to be avoided with relatively minimal risk of miscommunication, they may assign a disproportionately high cost to strategy sets that deviate too far from those rules.

5 Conclusions and future work

We conclude this paper with a brief discussion of areas in which our model goes beyond existing theories of both accent placement and game-theoretic pragmatics, some other phenomena where a similar approach would seem promising to us, and a suggestion of ways to test the predictions of the model.

5.1 Novel aspects of the model

Accent placement. As explained above, our model concerns the *rationalistic* factors that influence a speaker's choice of accent placement. It does not seek to address the rules of the grammar that relate accent placement

to information structure. However, some aspects of our proposal indirectly concern the architecture of that grammar. First, while we agree with the suggestion of German et al. (2006) and others that certain forms may be dispreferred despite being well-formed, we argue that such dispreferences should not be encoded in the grammar, but should be explained in terms of factors that lie outside of the grammar. Second, we argue that a mutually accessible probability distribution over possible speaker intentions (i.e., focus assignments) plays a key role in interlocutors' selection and interpretation of accent placement. Much previous work has neglected to reconcile the explicit assumption that speakers freely choose an intended focus structure with the implicit assumption that intuitions about felicity are a reflection of grammatical constraints. In our model, we straightforwardly adopt the first assumption and propose that intuitions about felicity may be explained by the fact that interpretation is guided by a mix of forces, including mutual beliefs and expectations. This offers a way to reconcile two widespread but seemingly contradictory assumptions in the theory of focus and accent placement.

Game-theoretic pragmatics. Game theory has been particularly useful for modeling the ways in which interlocutors enrich the conventional meaning of forms. In Parikh's (2001) analysis of scalar implicatures, for example, strategic inferences make it possible for an utterance of *'Some of the boys went to the party'* to convey the truth-conditionally stronger meaning of the sentence *'Some of the boys went to the party, and not all of the boys went to the party'*. Notice that the latter meaning entails the former. In fact, it is typical of game-theoretic analyses that the meanings at issue are monotonically related in some way.[13] By comparison, our analysis does not assume any particular relationship between the grammatically determined interpretation of an accent pattern and the interpretation that results from strategic inference. Our model does not require, in other words, that the pattern of F-marking in (10a) and those in (10b) and (10c)[14] be related in any particular way.

[13]Consider Parikh's (2001) analysis of relevance implicatures, for example, in which a sentence like *'It's 4pm'* is enriched with the meaning of *'Let's go for the talk'*. Even though the latter is not "semantically related" to the former in the same way that *'Some of the boys went to the party'* is related to *'Not all of the boys went to the party'* (Parikh, 2001, p. 93), the inferred content is monotonically added (via logical conjunction) to the conventional meaning of the utterance.

[14]Recall that, because of focus projection, accentuation on the direct object is grammatically consistent with multiple patterns of F-marking.

(10) a. ...they play their game $[\text{IN}]_F$
b. ...they play $[\text{their GAME}]_F$ in
c. ...they $[[[\text{play}]_F\ [\text{their GAME}]_F]_F\ [\text{in}]_F]_F$

Instead, the alternatives to the grammatically licensed interpretation arise merely because they correspond to different ways of assigning *F*-marking to the underlying syntactic representation. Certain alternatives then emerge as more relevant to the game structure because the context renders them more probable than others.

5.2 Extending the analysis

There are additional cases in which weak prosodic restrictions may be interacting with discourse-related constraints. The ones we address in this section differ from our own in a number of ways, yet we feel that there is an underlying similarity among them that warrants a common treatment.

Tone Compression in German. Languages vary in the way they treat complex intonation contours applied to monosyllables (Ladd, 1996, 132-4). In English, rise-fall-rise contours may be associated with a single syllable, as in (11):

(11) Sue?!
L+H* L-H%

English works differently in this respect from German. Examples (12) and (13) illustrate the high-fall-rise intonation contour that marks questions in German (Ladd, 1996, 133).

(12) Ist das Ihre TÜTE?
H* L-H%
'Is this your BAG?'

(13) #Ist das Ihr GELD?
H* L-H%
'Is that your MONEY?'

In (12), the three tones associated with this contour are realized over two syllables,[15] such that no syllable carries more than two tones. In (13), by comparison, all three tones are compressed onto a single syllable.

[15] Note that '*Tüte*' is pronounced /tyːˈtə/.

According to Ladd (1996), (13) has a "phonetically degraded" quality, even in contexts in which it is expected to be pragmatically appropriate (e.g., someone has left some money on the table). Ladd suggests that in such cases, a speaker is likely to substitute an alternative form, such as (14), in which the three tones are realized over at least two syllables.

(14) Ist das IHR Geld?
H* L-H%
'Is that YOUR money?'

Crucially, however, (13) and (14) do not seem to carry the same meaning. In terms of our earlier framework, (14) marks the expression *'Geld'* as given, and is predicted to be most appropriate when money has been explicitly mentioned in the discourse, while (13) is more appropriate otherwise. In short, the form-based preference for (14) over (13) appears to outweigh the speaker's desire to mark the status of *'Geld'* in the contextually most appropriate way.

This suggests an interesting twist on our analysis of stranded prepositions. In that analysis, we claimed that certain focus structures may be conveyed *in spite of* being inconsistent with what is required by focus projection rules. By comparison, it does not seem likely that (14) can be used to *convey* the meaning of (13). On the contrary, (14) intuitively seems to require that the listener accommodate the fact that *'Geld'* is given. This suggests that the relevant tradeoff is not between form-based costs and mismatch costs as in our earlier example, but between form-based costs and the utility that the speaker associates with each of the possible meanings. If the speaker assigns roughly equal utility to each way of assigning a status to *'Geld'*, for example, with perhaps a slight preference for treating *'Geld'* as non-given as in (13), then there is an increased potential for factors other than context to influence the speaker's choice. In this particular case, the preference for avoiding tone compression is sufficient to favor the pattern in (14). If, however, the speaker were to associate a much higher utility with treating *'Geld'* as non-given as compared with treating it as given, then the speaker will prefer (13), and any preference between the two forms is unlikely to affect his or her decision.[16]

[16]It is interesting to note that the *preferred* form in this case involves a nuclear accent on *'Ihr'* 'your', which is a possessive pronoun and therefore a function word. To the extent that German et al.'s findings for prepositions generalize to other function word categories, this would be somewhat unexpected. In the end, however, the preference for (14) is observed independently on the basis of phonological well-formedness (however impressionistic), so this does not pose a problem for our approach. It does suggest, however, that in this

Since we are assuming for the present case that the form-meaning association given by the grammar is fixed, there are effectively only two strategy sets to consider, one in which the speaker uses (13) to communicate (successfully) that *'Geld'* is non-given, and one in which the speaker uses (14) to communicate that *'Geld'* is given. In that sense, the problem may in fact reduce to a simple decision problem for the speaker between which of the two strategy sets has the higher expected utility. Nevertheless, problems like this have important aspects in common with our above analysis and suggest a certain generality to the approach we are taking.

Prosodic Promotion in English. Pitch accent assignment is often discussed in terms of its tendency to encode information structure. Across speakers, utterances, and even phrases within utterances, however, there is substantial variability in the density of accent assignment. Pierrehumbert (1994) proposes to account for this variability through a phonological process called *prosodic promotion*. Normally, pitch accents that do occur, tend to occur on prominent syllables. Prosodic promotion, however, allows syllables that would normally be too weak to carry a pitch accent to be "promoted" to a higher level of prominence so that they may carry one. This process, she argues, is not only sensitive to "discourse factors," but it is "generally available to strengthen prosodically weak elements if the speaker for any reason wishes to accent them."

Notice, however, that prosodic promotion may have consequences for the prosodic marking of information structure. In general, the distribution of accents in an utterance encodes information about the focus assignment because accents are (i) required to occur within a focus and (ii) excluded or at least minimized elsewhere (Schwarzschild, 1999; Truckenbrodt, 1995; Williams, 1997; Sauerland, 2005; Büring, 2008). In the most canonical examples, the focused constituent includes a nuclear accent close to its right edge, while the post-focal material is free of pitch accents or *deaccented* up to the end of the utterance. The result is a kind of discontinuity in the pitch accent distribution that allows the hearer to identify the rightward extent of the focus. When accents are freely assigned as a result of prosodic promotion, however, then the discontinuity may disappear, thereby obscuring the location of the right edge of the focus. Applied to our earlier example, the location of the last accent would normally distinguish the focus assignment in (10a) from those in (10b) and (10c).

---

particular context, the phonological preference against tone compression outweighs any similar preference against accenting possessive pronouns.

(10) a. ...they play their game $[\text{IN}]_F$
b. ...they play $[\text{their GAME}]_F$ in
c. ...they $[[[\text{play}]_F\ [\text{their GAME}]_F]_F\ [\text{in}]_F]_F$

If prosodic promotion applies at the level of the syllable, however, then all words, including *'game'* and *'in'*, will receive an accent as in (15), and all information about the focus assignment will be lost. In short, prosodic promotion may be detrimental to communication.

(15) THEY PLAY THEIR GAME IN

Consider also that accents are themselves associated with cost and effort, as is assumed, for example, by the wide range of proposals that link accentuation with markedness (Schwarzschild, 1999; Beaver, 2004; Clark and Parikh, 2007; Hirschberg and Ward, 1991). In addition, it is widely held that pitch accents represent a type of morpheme (Liberman and Sag, 1974; Liberman, 1975; Gussenhoven, 1983; Ladd, 1980; Pierrehumbert and Hirschberg, 1990; Pierrehumbert, 1994). On standard assumptions of both neo-Gricean and game-theoretic analysis, then, an increase in the number of accents would result directly in an increase in morpho-syntactic complexity, which in turn is associated with an increase in production effort and processing costs.

If prosodic promotion results in utterances that are not only more costly to produce, but potentially carry less information about the speaker's intention, then what factors would lead a speaker to use it? One factor may be the need to overcome situational impediments to the speech signal, as in cases of ambient noise or unreliable channels. In such a case, a speaker applies prosodic promotion as a way to increase not only the overall amplitude of the utterance, but the acoustic distinctness of individual syllables and phonemes. This tends to insure that the lexical content of the speaker's utterance is recovered, even if other aspects of its meaning are forfeited.

In terms of the model we have outlined, this suggests that in cases where prosodic promotion would obscure the relationship between accent placement and focus assignment, the speaker must decide whether the risk of miscommunicating the lexical content outweighs both (i) the risk of miscommunicating focus, and (ii) the extra cost incurred by any additional pitch accents. Generally, this will be a function of the probability of successful communication of the lexical content given the degree of the impediment, as well as the prior probability that is assigned to the speaker's intended focus assignment given the context. When the probability of unsuccessful communication of the lexical content is high, and the context includes strong cues or the intended focus structure, then the speaker is likely to accept

the added costs and risks associated with prosodic promotion. By contrast, when that probability is low, and the context includes only weak cues to focus structure, then the speaker will prefer to minimize accentuation so that focus structure is maximally encoded by the accent distribution.

5.3 Testing the model

Our model predicts that accent placement should be sensitive to how likely the various parts of an utterance are to be treated by the speaker as given. Throughout the literature on information structure and reference resolution, the tendency for an expression to be interpreted as given or new is linked to various psycho-attentional properties of its potential antecedents. Often, these are collapsed into a single notion of either salience or accessibility, though specific proposals suggest a relevant role for the recency of the antecedent (Arnold, 1998), the morphosyntactic status of the antecedent (Prince, 1981; Gundel et al., 1993; Grosz et al., 1995), or the embedding relationship between adjacent discourse segments (Grosz and Sidner, 1986; Nakatani, 1997).

In the context of our model, such factors are predicted to play an important role in the way that probability distributions are assigned to focus structures in the shared beliefs of the speaker and hearer. Moreover, to the extent that they can be manipulated in a concrete way, it should in principle be possible to test the predictions of the model under a varied set of conditions. If, for example, expressions with more recent antecedents are treated as given in those focus structures that have the highest prior probability, then our model predicts that accentuation of the preposition in examples like (4) will actually be *less* likely when the antecedent is very recent.

Notice that this presupposes that the speaker's intention is fixed across all contexts. For the German et al. study, it was assumed that the speaker's intended focus assignment was known precisely because the antecedent utterance was very recent. In effect, the subjects were playing the role of both speaker and hearer simultaneously. When the same type of reading task is used for a study which varies the recency of the antecedent, however, then the interpretation that the speaker intends to represent may actually covary with the factor being manipulated. In other words, it is no longer reasonable to assume that the speaker's intended focus assignment is fixed across contexts in such a case. Thus, a study like the one we describe will ultimately require a more sophisticated approach, specifically, one that controls both (a) the focus assignment that the speaker hopes to convey, and (b) the focus assignment that the speaker believes to have the highest prior probability

in the mind of the hearer.

Acknowledgments

We are grateful to the editors for their support and extreme patience. Parts of this work were presented at *Games and Decisions in Pragmatics II* in Berlin, Germany, 2006. We thank the audience there for helpful comments. SK gratefully acknowledges the generous support of the American Council of Learned Societies and the Lichtenberg-Kolleg at the University of Göttingen, Germany.

References

Anttila, A. 1997. *Variation in Finnish Phonology and Morphology*. PhD thesis, Stanford University.

Arnold, J. 1998. *Reference Form and Discourse Patterns*. PhD thesis, Stanford University.

Beaver, D. 2004. The optimization of discourse anaphora. *Linguistics and Philosophy*, 27:3–56.

Beckman, M. and J. Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook*, III:15–70.

Benz, A., G. Jäger, and R. van Rooij, editors. 2006. *Game Theory and Pragmatics*. Palgrave MacMillan.

Boersma, P. and B. Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry*, 32:45–86.

Büring, D. 2008. What's given (and what's new) in the theory of focus. In *34th Meeting of the Berkeley Linguistics Society*.

Chomsky, N. 1972. *Studies on Semantics in Generative Grammar*, chapter Deep Structure, Surface Structure, and Semantic Interpretation, pages 62–119. Mouton, The Hague.

Chomsky, N. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Foris Publications.

Clark, H. 1996. *Using Language*. Cambridge University Press, Cambridge, UK.

Clark, H. and D. Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition*, 22:1–39.

Clark, R. and P. Parikh. 2007. Game theory and discourse anaphora. *Journal of Logic, Language and Information*, 16:265–282.

Croft, W. 1990. *Typology and Universals*. Cambridge University Press, Cambridge, UK.

Dekker, P. and R. van Rooij. 2000. Bi-directional Optimality Theory: An application of Game Theory. *Journal of Semantics*, 17:217–242.

Eckman, F. R., E. A. Moravcsik, and J. R. Wirth, editors. 1986. *Markedness*. Plenum Press, New York.

German, J., J. Pierrehumbert, and S. Kaufmann. 2006. Evidence for phonological constraints on nuclear accent placement. *Language*, 82:151–168.

German, J. S. 2008. *Prosodic Strategies for Negotiating Reference in Discourse*. PhD thesis, Northwestern University.

Grice, H. P. 1989. *Studies in the Way of Words*. Harvard University Press, Cambridge, MA.

Grosz, B., A. Joshi, and S. Weinstein. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21: 203–226.

Grosz, B. and C. Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12:175–204.

Gundel, J., N. Hedberg, and R. Zacharsky. 1993. Cognitive status and the form of referring expressions in discourse. *Language*, 69:274–307.

Gussenhoven, C. 1983. Focus, mode and nucleus. *Journal of Linguistics*, 19:377–417.

Haiman, J. 1985. *Natural Syntax, Iconicity, and Erosion*. Cambridge University Press, Cambridge, UK.

Hirschberg, J. and G. Ward. 1991. Accent and bound anaphora. *Cognitive Linguistics*, 2:101–121.

Kadmon, N. 2001. *Formal Pragmatics: Semantics, Pragmatics, Presupposition and Focus*. Blackwell Publishers, Oxford.

Ladd, D. R. 1980. *The Structure of Intonational Meaning*. Indiana University Press, Bloomington.

Ladd, D. R. 1996. *Intonational Phonology*. Cambridge University Press.

Lewis, D. 1969. *Convention*. Harvard University Press, Cambridge, MA.

Liberman, M. 1975. *The Intonation System of English*. PhD thesis, MIT. Reprinted by Garland Press.

Liberman, M. and I. Sag. 1974. Prosodic form and discourse function. In *Papers from the 10th Regional Meeting of the Chicago Linguistic Society*, pages 416–426.

Nakatani, C. 1997. Discourse structural constraints on accent in narrative. In van Santen, J., R. Sproat, J. Olive, and J. Hirschberg, editors, *Progress in Speech Synthesis*, pages 139–156. Springer Verlag, New York.

Parikh, P. 2001. *The Use of Language*. CSLI Publications, Stanford, CA.

Parikh, P. 2006. Pragmatics and games of partial information. In Benz et al. (2006), pages 101–122.

Parikh, P. 2010. *Language and Equilibrium*. The MIT Press, Cambridge, MA.

Pierrehumbert, J. 1980. *The Phonology and Phonetics of English Intonation*. PhD thesis, MIT. Distributed 1988, Indiana University Linguistics Club.

Pierrehumbert, J. 1994. Alignment and prosodic heads. In *Proceedings of the Eastern States Conferences on Formal Linguistics 10*, pages 268–286.

Pierrehumbert, J. and J. Hirschberg. 1990. The meaning of intonation and the interpretation of discourse. In Cohen, P., J. Morgan, and M. Pollack, editors, *Intentions in Communication*, pages 271–311. The MIT Press.

Prince, E. 1981. Toward a taxonomy of given-new information. In Cole, P., editor, *Radical pragmatics*, pages 223–256. Academic Press, New York, NY.

Roberts, C. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. *OSU Working Papers in Linguistics*, 49. Papers in Semantics.

van Rooij, R. 2004. Signalling games select Horn strategies. *Linguistics and Philosophy*, 27:493–527.

Sauerland, U. 2005. Don't interpret focus! Why a presuppositional account of focus fails and how a presuppositional account of givenness works. In Maier, E., C. Bary, and J. Huitink, editors, *Proceedings of Sinn und Bedeutung 9*, pages 370–384.

Schmitz, H.-C. 2005. *Optimale Akzentuierung und Active Interpretation*. PhD thesis, University of Bonn.

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

Selkirk, E. 1995. The prosodic structure of function words. In Beckman, J., L. W. Dickey, and S. Urbanczyk, editors, *Papers in Optimality Theory*, pages 439–470. GLSA Publications, Amherst, MA.

Selkirk, E. 1996. Sentence prosody: Intonation, stress, and phrasing. In Goldsmith, J. A., editor, *The Handbook of Phonological Theory*, pages 550–569. Basil Blackwell, London.

Thompson, W. 2009. *A Game-Theoretic Model of Grounding for Referential Communication Tasks*. PhD thesis, Northwestern University.

Truckenbrodt, H. 1995. *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*. PhD thesis, MIT.

Williams, E. 1997. Blocking and anaphora. *Linguistic Inquiry*, 28:577–628.

Winkler, S. 1997. *Focus and Secondary Predication*. Mouton de Gruyter, Berlin.