

# Toward a formalisation of speech-act functions of questions in conversation

Marina Terkourafi <sup>a</sup> Aline Villavicencio <sup>b</sup>

<sup>a</sup>*The British School at Athens, 52, Odhos Souidhias, GR-10676 Athens, Greece.*

<sup>b</sup>*University of Cambridge Computer Laboratory, William Gates Building, JJ Thomson Avenue, Cambridge, CB3 0FD, UK.*

---

## Abstract

In this paper we address the relationship between questions as grammatical and semantic entities and questions as pragmatic entities, arguing that a contextually-conditioned association holds between the former, as interrogative formulae, and the latter, as particular types of speech acts (namely offers and requests). This argument is supported by evidence from a corpus of spontaneous Cypriot Greek exchanges. This is an important observation which may lead to a gain in both speed and naturalness if it is captured in the grammar. We investigate a possible approach for formalising this observation using HPSG, wherein morphosyntactic, semantic and pragmatic constraints can be jointly represented. More specifically, we build on the notion of defaults to suggest that a default specification of illocutionary force can be associated with each formula providing an indication for its interpretation. This specification may nevertheless be overridden by further processing of contextual features, including intonation.

---

## 1 Introduction

The oft-noted terminological multifunctionality of the term ‘question’ (e.g. Groenendijk and Stokhof, 1997; Higginbotham, 1996) has been explicated as referring to three distinct kinds of entities: a) a grammatical one (where grammar encompasses both syntax and phonology), i.e. sentences of interrogative form (henceforth Q1); b) a semantic one, i.e. a semantic object distinct from propositions which is the ‘content/sense’ of a question (henceforth Q2), and c) a pragmatic one, identified with the speech act of question (henceforth Q3). Distinguishing between these three entities allows us to accommodate their several permutations in natural language,

---

*Email addresses:* mt217@cam.ac.uk (Marina Terkourafi),  
Aline.Villavicencio@cl.cam.ac.uk (Aline Villavicencio).

thus providing a unified account of different types of questions (e.g. rhetorical, didactic, unanswerable, as well as the more prototypical ones).

This paper focuses on the relationship between questions in the Q1/Q2 sense with questions in the Q3 sense. We begin by drawing some distinctions: one between Q3-questions and requests, a further one between requests for verbal action and requests for nonverbal action, and a final one between offers and requests. We then suggest that which type of act a Q1/Q2 question realises in context can be predicted in a nonmonotonic fashion by jointly taking into account its lexico-grammatical features and the values of a fixed set of extra-linguistic features. This proposal is empirically grounded in the analysis of a corpus of spontaneous Cypriot Greek (hence CG) exchanges, and formalised in the framework of Head-driven Phrase Structure Grammar.

Similar to latest approaches such as Asher and Lascarides (2003:304ff.), the current proposal attempts to resolve the well-known non-alignment of linguistic form with illocutionary act not by an appeal to a formally non-tractable notion of speaker's intention, but by exploiting information already given in the discourse. Where our proposal differs from that of Asher and Lascarides's is, first, in using information attached to lexical entries to motivate defeasible inferences regarding the type of speech act performed (something which they achieve with recourse to knowledge of the discourse structure and rhetorical relations applicable therein); and second, in being grounded in corpus data, whereas their analysis is formally motivated. One advantage of having based our analysis on a corpus of spontaneous interactions is that the addressee's uptake (Austin, 1962:115-6) was available to us, and has been exploited in classifying acts under different types. In this way, our analysis is faithful to native speakers' intuitions, and can be expected to yield more accurate predictions. On the other hand, the analysis must be repeated anew for each language, thereby prioritising faithfulness over generality/universality.

## **2 Theoretical preliminaries: distinguishing questions, requests, and offers**

On the traditional view, which classifies questions (in the Q3 sense) as directives ('requests for information'; (Searle, 1996[1979]:147-8; Bach and Harnish, 1979:47-8; Vanderveken, 1990:11), the illocutionary act realised by a question may be paraphrased as: "I hereby ask you to answer (the question) Q", or "Please tell me/bring it about that I know Q" (with Q same as before). One may remain sceptical about such a move. For one, questions have answerhood conditions, while requests (for verbal or nonverbal action) have compliance conditions (Harnish, 1994:415; Vanparys, 1996:35). Related to this, a question may be posed without the expectation of an answer, or, for that matter, any other action in return, as is the case with rhetorical and unanswerable questions (Lyons, 1995:452; Récanati, 1987:158; Sperber and Wilson, 1995[1986]:249-54; Vanparys, 1996:73).

Alternatively, questions may be analysed as a class of speech acts in their own right, i.e. separate from directives. Paraphrasing Lyons (1977:756), we may explicate the illocutionary act performed in this case as “posing” a question, i.e. expressing doubt/lack of ability to assign a truth value to a particular state of affairs (or any one out of a set of related states of affairs, in the case of *wh*-questions). The intention for the hearer to respond by providing the ‘missing’ truth value is not part of performing the speech act of question as such (Lyons, 1977:755). However, the intention to get the hearer to recognise that the speaker has a particular proposition (better: propositional concept; Groenendijk and Stokhof, 1997:13-4) in mind and that *s/he* is entertaining it in the dubitative mode is (Lyons, 1977:756). Or, in Groenendijk and Stokhof’s words (1997:18), “interrogative acts [our Q3; MT and AV] do not necessarily direct the hearer towards any kind of action, except perhaps that of considering the question.”

Adopting the latter perspective, it is possible to distinguish questions, which bear no requirement to be provided with an answer, from requests (which do bear a requirement to be complied with). Thus, an utterance from the CG data such as:

(1) [On the radio; Speaker: female, aged 31-50, middle-class; Addressee: male, aged 31-50, middle-class; Relationship: interviewer to interviewee]

<i>prin</i>	<i>arçisume</i>	<i>omos</i>	<i>mas lete</i>	<i>parakalume poli</i>
before	we start	however	would you tell us	please
<i>ti</i>	<i>itan</i>	<i>afti</i>	<i>i kaskarika tis: grinpis ...</i>	
what	was	this	Greenpeace blunder...	

constitutes a request for verbal action (“would you please tell us”). Comparing this with the corresponding question:

(2) *ti*      *itan*      *afti*      *i kaskarika tis: grinpis ...*  
 what      was      this      Greenpeace blunder...

we see that (1), but not (2), allows significant scope for encoding features pertaining to the interactional dimension of the discourse, and may, in virtue of this potential, be used by speakers with an interest in encoding such features in the form of their utterance – which is indeed what we find in the CG data, where information-seeking utterances in formal settings (on radio/TV, in formal discussions) routinely take elaborate forms such as (1), as opposed to (2), which in turn prevails in informal settings (at home, at work). Speakers’ choice between a question and a request for verbal action would, then, seem to be significant, such that the two cannot be conflated without some loss in descriptive accuracy.

The distinction between requests for verbal action and requests for nonverbal action, on the other hand, is grounded in a very general distinction between information-focussing and action-focussing utterances, which may be explicated by appealing

to the notion of ‘direction of fit’ (Searle, 1996:142; Harnish, 1994:415):

*“...the world-to-words fit that is the illocutionary point of an interrogative act [our request for verbal action; MT and AV] seems to be of a different kind than that of a simple request. The latter calls for an action that transforms the world as such, whereas an informative question ... demands that a change be brought about in the information state of the speaker ... Of course, any utterance transforms the world in this, admittedly rather trivial way ... But for these kinds of speech acts that is not the real point: they require a change in information about the world, not in the world itself.”* (Groenendijk and Stokhof, 1997:17)

In line with these insights, we claim that what binds together our two subtypes of request is their common world-to-words direction of fit. Questions, on the other hand, exhibit the words-to-world direction of fit, in that, in posing a question, the speaker makes known his/her doubt (i.e. describes a situation whereby s/he is in doubt) regarding a particular state of affairs (or set of related states of affairs). At the same time, the world-to-words fit of requests for verbal action is essentially distinct from that of requests for nonverbal action along the lines described above, which justifies keeping apart the two subtypes of requests.

Our final distinction, between offers and requests, is traditionally formulated as follows (Searle, 1996:147-8; Bach and Harnish, 1979:47-51): directives express the attitude of the speaker towards some future act of the addressee, along with his/her intention that his/her utterance be taken as reason for the addressee to act, while commissives express the speaker’s commitment to a future course of action, specified in the propositional content of his/her utterance, along with the intention that his/her utterance be taken as reason for the addressee to believe that s/he commits him/herself to this course of action. While the traditional definition’s reference to the propositional content of the utterance (an echo of Searle’s 1969:63 ‘propositional content’ condition), is not always corroborated by the data, we nevertheless argue for a renewed appreciation of the importance of propositional content, namely as indicating to whom (the speaker or the addressee) the state of affairs referred to is desirable. The illocutionary force inferred based on the ‘locus of desirability’ of the speech act may on occasion be different from the illocutionary force inferred based on which of the two participants is responsible for bringing about the state of affairs referred to. In such cases, examples show that the illocutionary force inferred based on which of the two participants constitutes the locus of desirability yields the intuitively correct result, and should therefore take precedence. Distinguishing between offers and requests along a general dimension of desirability (to the speaker or to the addressee) is furthermore empirically supported by evidence that speech-act theoretic descriptions play no actual role in comprehension (Good MS, reported in Geis, 1995:31) and illustrated with examples from CG.

Summing up, further to questions, which bear no expectation of an answer, two classes of performative acts may be distinguished, a generic class of requests, referring to states of affairs desirable from the speaker’s point of view, and a generic

class of offers, referring to states of affairs desirable from the addressee's point of view. Usually, but not always, the responsibility for bringing about the state of affairs referred to lies with the addressee for requests, and with the speaker for offers. At a finer level of detail, requests for verbal action can be distinguished from requests for nonverbal action based on the difference in how the world is to come to fit the words in each case (through a change in information about the world vs. a change in the world), and the variability in degree of imposition carried by each (cf. Brown and Levinson, 1987:77). In this paper we shall be focussing on questions in the Q1/Q2 sense used to perform requests for nonverbal action and offers.

### 3 The Cypriot Greek corpus data

With these preliminaries in place, we may now turn to an examination of the CG conversational data. These originate in approximately 110 hours of recordings of spontaneous face-to-face exchanges between adult native CG speakers of both sexes and various socio-economic backgrounds conducted in the four major urban centres of the island, and in a variety of settings, generally classified as: at home/informal social gatherings; at work (mainly shops and offices); on radio/TV/at formal discussions. During the collection of the data, information pertaining to the following situational features was noted: the interlocutors' sex, approximate age, and social class, the relationship between them, the setting of the exchange, and whether the speech act performed occurred for the first or subsequent times. Frequency counts revealed that particular combinations of values of extra-linguistic variables tend to co-occur with particular expressions, which we may, for our purposes, conceptualise as 'formulae' (cf. Pawley and Syder's, 1983:208ff. 'lexicalised sentence stems'), that is, morpho-phonologically specified surface realisations of VPs, where the head of the VP is subject to lexical/semantic constraints.

In this paper, we discuss three formulae which provide the nucleus for interrogative sentences in the data: *eçete NP?* (**do-you.pl-have NP?**), *efi NP?* (**does-it-have NP?**), and *θelis NP/VP?* (**do-you.sg.-want NP/to VP?**). Our argument is that, in the CG data, there is an association between the interrogative sentences instantiating these formulae and the type of act performed each time, but this association only holds in particular contexts of use (combinations of values of the extra-linguistic features), and then only in a probabilistic fashion.

Thus, *eçete NP?*, as in (3), prevails in opening requests from middle-class customers previously unacquainted with the service-provider facing them. In the corpus, out of 9 occurrences of *eçete NP?*, 6 fell into this characterisation; another was used by a customer to address a familiar service-provider, while the remaining 2 realised offers. Moreover, *eçete NP?* is the most preferred formula used by customers in these circumstances, being used 15% of the time, with other verb-forms following at 12% of the time or less.

(3) [In a pub; Speaker: male, aged 31-50, middle-class; Addressee: female, aged 18-30, working-class; Relationship: new customer to salesperson]

*eçete*                      *p<sup>h</sup>inats?*  
*Do you-2pl. have*      some peanuts?

*efi NP?* fulfils a similar role as *eçete NP?*, but for working class customers, as in (4):

(4) [In the open-air market; Speaker: female, aged 31-50, working-class; Addressee: male, aged 31-50, working-class; Relationship: new customer to salesperson]

*efi*                      *mikres pu na min exun sçeđia pano? aspro*  
*Are there*              any small plain ones?                      In white.

In the corpus, out of 23 occurrences of *efi NP?*, 20 are compatible with this characterisation, while the remaining 3 also realise requests but not opening ones. Moreover, *efi NP?* is the most preferred formula used by working-class customers facing unfamiliar service-providers, being used 14.8% of the time (full details of the frequencies for these formulae are given in Terkourafi, 2002).

As can be seen, the proposed association between the interrogative sentences instantiating these formulae and the type of act performed each time has two determinants: frequency of use of the formula to perform the particular act in a certain context, and range of contexts across which the formula is the most frequent expression used to perform the particular act. High scores on both determinants mean a stronger association, which in turn starts to become ‘visible’ by manifesting itself also structurally. This is what happens in the case of our last formula, *θelis NP/VP?*, as in (5) and (6) below (cf. Terkourafi, forthcoming):

(5) [In a shoe-shop; Speaker: female, aged 18-30, working-class; Addressee: female, aged 18-30, middle class; Relationship: acquaintances]

*'lis*                      *kafe? (.)*      *indalos*      *in'*                      *o kafes su?*  
*Do you want*      coffee?      how      do you take      your coffee?

(6) [In a shoe-shop; Speaker: female, aged 31-50, working-class; Addressee: female, aged 31-50, working class; Relationship: salesperson to new customer]

*θelis*                      *na valumen*      *kanena pataci*      *mesa?*  
*Do you want us*      to put                      an insole                      in?'

This formula is preferred to perform offers over a wide range of contexts (various combinations of speaker and addressee at home and at work, crucially including contexts where the sum of the social Distance and relative Power between interlocutors (Brown and Levinson, 1987: 76) are both high and low; Terkourafi, 2001:

95ff.), in each of which it constitutes the preferred (most frequent) realisation. This increased contextual autonomy for *θelis NP/VP?* goes hand in hand with a high degree of fixedness on both the paradigmatic and syntagmatic axes.

According to Lehmann (1985), the freedom with which a linguistic sign is used may be measured on three dimensions: weight (the degree to which a sign is distinct from members of its class), cohesion (the degree to which a sign systematically contracts certain relations with other signs), and variability (the degree to which a sign enjoys momentary mobility/shiftability with respect to other signs), each of which has a syntagmatic and a paradigmatic aspect. A loss in weight and variability, and/or an increase in cohesion translate into a gradual loss of freedom for the sign.

The loss in paradigmatic weight (or ‘integrity’) for *θelis NP/VP?* concerns two levels. At the level of form, its surface realisation, *θelis*, can be reduced to monosyllabic *θes?*, *’lis?* or *θe*; 18 out of 106 offers (17%) using *θelis?* are thus truncated, with *θes?* accounting for 15 (84%) of these. At the level of content, 15 out of 33 times (45.5%) when an uptake is provided, this consists in thanking or providing a request, showing that *θelis?* is not so much interpreted as a literal question about the addressee’s desires but as a token of the speaker’s availability/willingness to satisfy these, i.e. as an offer. Characteristically, a statement of the addressee’s desires using *θelo*, **I-want** is never provided in response to *θelis?*. On the other hand, the fact that *θelis?* cannot be adverbially modified (e.g. by an adverb denoting degrees of desire) testifies to a loss in syntagmatic weight. An increase in paradigmatic cohesion also appears to be taking place: of 129 occurrences of *θelis?* at home and at work, 106 (82.17%) perform offers, while only 23 (17.82%) perform requests. That is, *θelis?* is typically associated with the pragmatic paradigm of the offer. Moreover, the lexeme *θelo* occurs in the indicative with rising intonation-2sg. in 129 out of 201 (64.2%) occurrences at home and at work, i.e. a high degree of syntagmatic cohesion appears to exist between these particular values for main-clause verb-subjective modality (which encompasses intonation)-number/person. In addition, a loss in paradigmatic variability means that desire in offers at home and at work is overwhelmingly expressed by using the lexeme *θelo*, (118 occurrences, 96%; 106 of these (90%) are in the indicative with rising intonation-2sg.) rather than other semantically equivalent expressions (5 occurrences, 4%). Finally, a loss in syntagmatic variability means that 98 times out of 106 (92.5%) *θelis?* occurs utterance-initially, the only items that can precede it being address terms, or the conjunction *lipon*, ‘so’.

The above analysis provides a strong indication that *θelis NP/VP?* used as a formula is associated with the illocutionary force of an offer. That is, an intimate connection links together the syntactic, semantic and pragmatic features of this expression, making possible a defeasible inference from the former to the latter. Our proposal is then to exploit this empirical finding by capturing this connection in the grammar.

#### 4 Toward a formalisation

To capture the connection between syntactic, semantic and pragmatic features of the formulae discussed in this paper, it would of course be possible, as noted by Paolillo (2000:255), to define a hybrid account, using one theory to express grammatical constraints (e.g. Government and Binding) and augmenting it with another to express pragmatic ones (e.g. Discourse Representation Theory). However, there is no need for separate theories to express these different types of conditions, if we use a sign-based formalism such as Head-Driven Phrase Structure Grammar (HPSG; Pollard and Sag, 1987,1994) which already integrates them in a unified account (for some proposals on how pragmatic constraints on linguistic forms can be captured in HPSG, see Green, 1996; Paolillo, 2000; Siegel, 2000; Bender, 2001). Taking this as a starting point, in this section we discuss some of the requirements that such a formalism must meet in order to model the phenomena addressed in this paper.

HPSG is a lexicalised formalism in which two main features are represented: PHON encodes the sign's sound content and SYNSEM captures syntactic, semantic and pragmatic information pertaining to the sign. Since the CG formulae discussed in this paper are obtained via different phonological rules (Standard Greek or Cypriot Greek ones) and this variation is socially significant (cf. Terkourafi, 2002), an IPA transcription of the phonological form of these formulae needs to be given under PHON. In addition, rising intonation needs to be indicated (as '?'), since interrogativity in Modern Greek is signalled prosodically rather than structurally (Sifianou, 1992:137-9). SYNSEM is divided into three attributes: CATEGORY includes the syntactic category of the sign, as well as its grammatical arguments; CONTENT captures its contribution to semantic (context-independent) aspects of the interpretation of any phrase that contains it; finally, CONTEXT represents pragmatic (context-dependent) information guiding use of the sign.

It is in CATEGORY that the morphosyntactic characteristics of a formula are defined, such as the fact that *θelis* is a second person singular verb form in the present tense. CONTENT is used to specify the semantics of the formula. This is where, among other things, the mode of the sentence (e.g. interrogative, imperative, etc) is specified, and where the particular speech act conveyed (e.g. offer, request, etc) can also be defined. Encoding the speech act as part of the formula's specification allows us to model the close association observed above between the formula and the speech act performed. However, as this is not always the case, a default specification needs to be employed. Thus, a particular formula realises by default a certain speech act only if certain pragmatic constraints are observed, such that the speech act specification is confirmed (or not) by the extra-linguistic context in which the formula occurs.

Defaults have been used to provide adequate descriptions for a range of natural

language phenomena (Daelemans et al., 1992; Lascarides and Copestake, 1999). Our use of defaults is in line with Lascarides and Copestake's (1999) proposal of persistent defaults at the interface between the lexicon and pragmatics, where lexical information is drawn upon to fix the semantics of certain words *ceteris paribus*; e.g., in the absence of pragmatic information to the contrary an intransitive usage of *drink* implies the drinking of an alcoholic beverage as in the sentence *Bill always drinks at parties*. As it is the pragmatic context which determines whether the default holds or not, persistent defaults are employed so that the defaults are not made into infeasible constraints after the lexicon, but are kept defeasible for further processing. For instance, if *Bill* is a five year old child who likes to drink orange juice, this default is overridden. Building on this tradition we argue for the use of persistent defaults to specify that a given formula tends to realise a certain speech act. For our purposes, defaults are used to indicate a bias in the production/interpretation of a formula, but can be overridden in case further processing disconfirms this bias.

The second attribute of SYNSEM, CONTEXT, has two attributes. The first one, contextual-indices, specifies the contextual anchors for an utterance: the speaker (SPKR), the addressee (ADDR), the time and location of the utterance, and so on. The second one, background (BKGD), includes assumptions about social aspects, and is typically about entities and events mentioned in the description of the semantic content of the sign. Background information is a set of propositions known as parameterised states of affairs (psoas), which can be thought of as felicity conditions on the utterance context. That is, if an utterance is not compatible with these psoas, it does not mean that the utterance is ungrammatical, but only that the context in which it is being used is not the expected one.

This general HPSG framework, or variations thereof, have been previously used to analyse pragmatic phenomena. For instance, Fukushima (2002) proposes an analysis of subject honorification in Japanese, where some affixes of the verb add a politeness component to the verb's background, indicating that its subject is *honoured*. Given that honouring relations can be specified among the pragmatic aspects of an NP, a sentence is pragmatically acceptable when the verb and the subject NP have compatible values for these constraints. A sentence where the subject is not compatible with these (e.g. where the subject is a *delinquent boy*, who is not usually honoured) will not be pragmatically valid, even though it can be syntactically and semantically acceptable.

For our purposes, information about the gender (PARTS-GENDER), age (PARTS-AGE), and social class of the speaker and addressee (PARTS-CLASS) needs to be specified, as it may bear on the selection of the formula to be used in a particular situation. In addition, it is necessary to capture the participants' social roles in the situation at hand (SOC-ROLE: e.g. friends, family members, colleagues, etc), the setting (SETTING) in which the exchange is taking place, and whether the act occurs for the first time or is repeated following an initial occurrence (SEQ-PLACE).

Different combinations of these features may determine the choice of formula. In HPSG such information can be encoded as BKGD constraints. A sign containing these attributes is shown in the following figure, in an Attribute-Value Matrix format (AVM):

SYNSEM:CONTEXT:BKGD :	<table style="border-collapse: collapse;"> <tr> <td style="padding: 2px;">PARTS-CLASS :</td> <td style="padding: 2px;">[</td> <td style="padding: 2px;">SPKR:CLASS : <b>class</b></td> <td style="padding: 2px;">]</td> </tr> <tr> <td style="padding: 2px;"></td> <td style="padding: 2px;"></td> <td style="padding: 2px;">ADDR:CLASS : <b>class</b></td> <td style="padding: 2px;"></td> </tr> <tr> <td style="padding: 2px;">SEQ-PLACE :</td> <td style="padding: 2px;"></td> <td style="padding: 2px;"><b>seq_place</b></td> <td style="padding: 2px;"></td> </tr> <tr> <td style="padding: 2px;">AGE :</td> <td style="padding: 2px;"></td> <td style="padding: 2px;"><b>relative_age</b></td> <td style="padding: 2px;"></td> </tr> <tr> <td style="padding: 2px;">PARTS-GENDER :</td> <td style="padding: 2px;">[</td> <td style="padding: 2px;">SPKR:GENDER : <b>gender</b></td> <td style="padding: 2px;">]</td> </tr> <tr> <td style="padding: 2px;"></td> <td style="padding: 2px;"></td> <td style="padding: 2px;">ADDR:GENDER : <b>gender</b></td> <td style="padding: 2px;"></td> </tr> <tr> <td style="padding: 2px;">SOC-ROLE :</td> <td style="padding: 2px;">[</td> <td style="padding: 2px;">SPKR:SOC-ROLE : <b>social-relation</b></td> <td style="padding: 2px;">]</td> </tr> <tr> <td style="padding: 2px;"></td> <td style="padding: 2px;"></td> <td style="padding: 2px;">ADDR:SOC-ROLE : <b>social-relation</b></td> <td style="padding: 2px;"></td> </tr> <tr> <td style="padding: 2px;">SETTING :</td> <td style="padding: 2px;"></td> <td style="padding: 2px;"><b>setting</b></td> <td style="padding: 2px;"></td> </tr> </table>	PARTS-CLASS :	[	SPKR:CLASS : <b>class</b>	]			ADDR:CLASS : <b>class</b>		SEQ-PLACE :		<b>seq_place</b>		AGE :		<b>relative_age</b>		PARTS-GENDER :	[	SPKR:GENDER : <b>gender</b>	]			ADDR:GENDER : <b>gender</b>		SOC-ROLE :	[	SPKR:SOC-ROLE : <b>social-relation</b>	]			ADDR:SOC-ROLE : <b>social-relation</b>		SETTING :		<b>setting</b>		
PARTS-CLASS :	[	SPKR:CLASS : <b>class</b>	]																																			
		ADDR:CLASS : <b>class</b>																																				
SEQ-PLACE :		<b>seq_place</b>																																				
AGE :		<b>relative_age</b>																																				
PARTS-GENDER :	[	SPKR:GENDER : <b>gender</b>	]																																			
		ADDR:GENDER : <b>gender</b>																																				
SOC-ROLE :	[	SPKR:SOC-ROLE : <b>social-relation</b>	]																																			
		ADDR:SOC-ROLE : <b>social-relation</b>																																				
SETTING :		<b>setting</b>																																				

The default specification of the speech act performed, as well as the contextual constraints governing the application of this default can be added to the formula concerned each time by means of lexical rules. Lexical rules are mechanisms which, when applied to an input sign representing a specific set of constraints, generate an output sign with a derived set of constraints. When a lexical rule is applied to produce a finite verb form (such as our CG formulae) from a base form, it brings with it a specification of the relevant contextual constraints under which the finite form realises the default speech act associated with it. One possibility would then be to define a family of lexical rules which would encode different combinations of contextual constraints for each formula applying at different levels of generality (depending on the contextual range of application of the formula).

## 5 Conclusion and Future Work

In this paper we present evidence from Cypriot Greek showing that a contextually-conditioned association holds between interrogative formulae (questions in the Q1/Q2 sense) and particular types of speech acts, namely offers and requests. If this is indeed so, a gain in both speed and naturalness lies in capturing this association in the grammar. To this end, we discuss a possible approach for formalising these observations, using HPSG. We address certain characteristics that make an HPSG analysis attractive, wherein the relevant contextual constraints could be specified via a lexical rule which produces the finite form of a verb from the base form simultaneously associating this with a default illocutionary force. This illocutionary force merely provides an indication as to the most likely interpretation of a formula but can be overridden by other features of the context incompatible with this interpretation including intonation. There are still open issues, such as the possibility of defining the probabilities associated with the formulae, and how this could be mapped into an HPSG formulation. Nevertheless, this is a promising approach which we intend to test in future work by investigating its applicability to

particular feature structures extrapolated from the Cypriot Greek data.

## References

- [Asher and Lascarides 2003] Asher, N. and Lascarides, A. *Logics of conversation*. Cambridge: Cambridge University Press, 2003.
- [Austin 1962] Austin, J.L. *How to do things with words*. Oxford: Clarendon Press, 1962.
- [Bach and Harnish 1979] Bach, K. and Harnish, R. *Linguistic communication and speech acts*. Cambridge, MA: MIT Press, 1979.
- [Bender 2001] Bender, E. *Syntactic variation and linguistic competence: the case of AAVE copula absence*. Ph.D. thesis, Stanford University, 2001.
- [Brown and Levinson 1987] Brown, P. and Levinson, S.C. *Politeness: some universals in language usage*. Cambridge: Cambridge University Press, 1987. Reprinted from: Goody, E. (ed.) *Questions and politeness: strategies in social interaction*. Cambridge: Cambridge University Press, 56-324, 1978.
- [Daelemans et al. 1992] Daelemans, W., De Smedt, K. and Gazdar, G. Inheritance in Natural Language Processing. *Computational Linguistics*, 18:2, 205-218, 1992.
- [Fukushima 2002] Fukushima, K. Competence and performance revisited: the implications of social role terms in Japanese. *Journal of Pragmatics*, 34, 938-968, 2002.
- [Geis 1995] Geis, M. *Speech acts and conversational interaction*. Cambridge: Cambridge University Press, 1995.
- [Green 1996] Green, G. M. The Structure of CONTEXT: The Representation of Pragmatic Restrictions in HPSG. In: Yoon, J. (ed.) *Studies in the Linguistic Sciences: Proceedings of the 5th Annual Meeting of the Formal Linguistics Society of the Midwest*, 1996.
- [Groenendijk and Stokhof 1997] Groenendijk, J. and Stokhof, M. Questions. In: Van Benthem, J. and Ter Meulen, A. (eds.) *Handbook of logic and language*. Oxford: Elsevier Science. 1055-1124, 1997.
- [Harnish 1994] Harnish, R. Mood, meaning and speech acts. In: Tsohatzidis, S. (ed.) *Foundations of speech act theory: philosophical and linguistic perspectives*. London: Routledge. 407-459, 1994.
- [Higginbotham 1996] Higginbotham, J. The semantics of questions. In: Lappin, S. (ed.) *The handbook of contemporary semantic theory*. Oxford: Blackwell. 361-383, 1996.
- [Lascarides and Copestake 1999] Lascarides, A. and Copestake, A. Default Representation in Constraint-based Frameworks. *Computational Linguistics*, 25:1, 55-105, 1999.
- [Lehmann 1985] Lehmann, C. Grammaticalisation: synchronic variation and diachronic change. *Lingua e Stile*, 20, 303-318, 1985.
- [Lyons 1977] Lyons, J. *Semantics*. Vol. II. Cambridge: Cambridge University

- Press, 1977.
- [Lyons 1995] Lyons, J. *Linguistic semantics*. Cambridge: Cambridge University Press, 1995.
- [Paolillo 2000] Paolillo, J.C. Formalizing formality: An analysis of register variation in Sinhala. *Journal of Linguistics*, 36: 215-259, 2000.
- [Pollard and Sag 1987] Pollard, C. and Sag, I.A. *Information-Based Syntax and Semantics*. CSLI Lecture Notes Series, n. 13, 1987.
- [Pollard and Sag 1994] Pollard, C. and Sag, I.A. *Head-driven Phrase Structure Grammar*. Chicago University Press, 1994.
- [Récanati 1987] Récanati, F. *Meaning and force*. Cambridge: Cambridge University Press, 1987.
- [Searle 1996] Searle, J. A taxonomy of illocutionary acts. In: Martinich, A.P. (ed.) *The Philosophy of Language*. Oxford: Oxford University Press, 141-155, 1996. Reprinted from: Searle, J., *Expression and meaning*. Cambridge: Cambridge University Press, 1-29, 1979.
- [Siegel 2000] Siegel, M. Japanese Honorification in an HPSG framework. In Ikeya, A. and Kawamori, M. (eds.) *Proceedings of the 14th Pacific Asia Conference on Language, Information and Computation*. Waseda University International Conference Center, Tokyo. Logico-Linguistic Society of Japan, 289-300, 2000.
- [Sifianou 1992] Sifianou, M. *Politeness phenomena in England and Greece: A cross-cultural perspective*. Oxford: Clarendon Press, 1992.
- [Sperber and Wilson 1995] Sperber, D. and Wilson, D. *Relevance: communication and cognition*. Oxford: Blackwell. Second edition, 1995[1986].
- [Terkourafi 2002] Terkourafi, M. Politeness and formulaicity: evidence from Cypriot Greek. *Journal of Greek Linguistics*, 3, 179-201, 2002.
- [Terkourafi forthcoming] Terkourafi, M. Pragmatic correlates of frequency of use: the case for a notion of 'minimal context'. In: Marmaridou, S. et al. (eds.) *Re-viewing Linguistic Thought: Perspectives into the 21st century*. Amsterdam: John Benjamins, forthcoming.
- [Vanderveken 1990] Vanderveken, D. *Meaning and speech acts*. Vol. II. Cambridge: Cambridge University Press, 1990.
- [Vanparys 1996] Vanparys, J. *Categories and complements of illocutionary verbs in a cognitive perspective*. Frankfurt: Peter Lang, 1996.