

# On the segmentation of requests in spoken language

Michael Alvarez-Pereyre

Université Paris-Sorbonne – CELTA

Maison de la recherche (D.310)

28 rue Serpente, 75006 Paris, France

maponline@gmail.com

## Abstract

This paper presents a theoretical approach to the characterization of requests boundaries and structure in general spoken dialogue. Emphasis is laid on the fracture between the illocutionary act of requesting (for which the term ‘request’ is kept) and the locutionary elements that carry it out (its ‘instantiation’). This approach leads to a representation of requests based on the inclusion of a semantic level under a pragmatic level via a structural level. These distinctions are meant to benefit to the semantic-pragmatic segmentation of dialogue and the study of request strategies.

## 1 Introduction

This paper focuses on the segmentation of requests in spoken language from a semantic-pragmatic perspective. Taxonomies exist for specific types of requests,<sup>1</sup> and general dialogue-acts taxonomies like DAMSL cover various types of utterances ‘influencing the addressee’s future action’ (Core and Allen, 1997; Stolcke et al., 2000). The aim is not to replace them—the former are finer-grained than what is proposed here, and the latter have the advantage of treating requests in a framework which includes other types of dialogue acts. The purpose of this paper is rather to contribute to a middle-ground, with distinctions general enough to encompass all types of action requests (assuming that a common process of ‘requesting’ underlies them), yet detailed and structured enough to account for the construction of their meaning. The goal is therefore not to provide a taxonomy *identifying* speech or dialogue acts in ‘shallow’ dis-

<sup>1</sup>For instance clarification requests or ‘CRs’ (Corsaro, 1977; Purver, Ginzburg and Healey, 2003; Purver, 2004; Rodríguez and Schlangen, 2004, among others), check-questions (Jurafsky and Martin, 2000, §19.3), etc.

course structure (Jurafsky et al., 1997), but a structured *explanatory* taxonomy of requesting means.

The scope is ‘requests for action’ taken in a broad sense, as exemplified by (1-3).<sup>2</sup>

- (1) Err / hmm / you know / it would probably be easiest if I just squeezed back there and poked around myself / would that be alright with you? // (BRO.0h32m56s)
- (2) Mister Masry? // [-Yeah //] I was wondering can you tell me who I talk to / about maybe getting an advance on my paycheck // Just / for the week-end // (BRO.0h14m33s)
- (3) Now you listen // I don’t give a damn / which way you go / just don’t follow me / you got that? // (FUG.0h18m59s)

‘Request’ is understood broadly to include the whole spectrum of invitations, entreaties, commands, etc. ‘Action’ is understood broadly in the sense that the scope includes requests for clarification (e.g. ‘Who said it?’); for attention, as ‘Now you listen’ in (3) or ‘Mr Masry?’ in (2); for confirmation, as ‘You got that?’ in (3); and of course what corresponds to a narrow understanding of the expression ‘action requests’, namely requests for actions not concerned with dialogue management, as the request to allow the speaker of (1) into the file room of the county water board, the request to direct the speaker of (2) to the right person, or the request not to follow the speaker of (3). The scope excludes ‘true’ questions (unmarked information

<sup>2</sup>The sequences quoted in this paper are extracted from a corpus of contemporary North-American films. Though film dialogues can by no means be called ‘spontaneous’ speech, they share enough features with naturally occurring interactions as to help us define the tools to study requests in spoken language. The advantage of working with commercial films is that such material covers the whole gamut of pragmatic interactions and situations—though, admittedly, as represented not ‘intercepted’ scenes. Sequences are indexed with three block capitals to identify the film quoted (e.g. *Erin Brockovich*, found at [BRO] in the References) and three numbers specifying the hour, minute and second when the sequence begins. The sound track is transcribed as the succession of speech ‘increments’ separated by pauses, with simple slashes [/] and double slashes [//] to distinguish between ‘tentative’ and ‘final’ pauses, following Pike (1945).

requests in which no significant attempt to further influence the addressee is traced).

The main question of this paper is: What definition of a ‘request’ should we work with, if we are to describe the outer boundaries and the inner complexity of requests in a way that enables modeling and quantification of request strategies?<sup>3</sup>

Section 2 starts from the difficulty of assigning boundaries to spoken language ‘requests’ in the framework of traditional speech act theory, and stresses key principles for pragmatic research based on spoken corpora. Section 3 proposes a minimal set of distinctions necessary to account for the inner organization of request instantiations. Section 4 assesses the approach.

## 2 Towards a definition of ‘requests’

One aspect of language interactions which tends to be oversimplified is the relationship between a ‘request’ and its instantiation.<sup>4</sup> True, with ‘indirect requests’, traditional pragmatics cast light on the gap between the ‘illocutionary act’ of requesting and the ‘locutionary act’ (or the ‘literal’ elements) used to express it. To bridge this gap, it focused on the contextual felicity-conditions of utterances, or on the conversational implicatures, the maxims, the inference rules or the cognitive faculties that enable us to construe their meaning (Austin, 1962; Searle, 1969; Searle, 1979; Grice, 1975; Perrault and Allen, 1980; Lenci, 1994, etc.); but all too often, the very examples given as a starting point to such analyses are far too simple, as the signifier of the request is almost invariably composed of one isolated, syntactically pure segment.<sup>5</sup> Pragmatic analyses of this kind encourage an idealized vision of language interactions, in which a request (and more generally a speech act) coincides perfectly with one stand-alone, clear-cut and atemporal piece of language.

More recent ‘cue-based’ (Jurafsky and Martin,

<sup>3</sup>This paper deals primarily with the theoretical foundations of methodology and does not tackle the technical implementation of the results.

<sup>4</sup>‘Instantiation’ might be preferred to ‘formulation’, as the former term makes it clearer that the speech elements uttered participate not so much in the communication as in the *performing* of the request (along with other elements not discussed here such as intonation, gesture, social context, etc.).

<sup>5</sup>Stubbs (1983, p. 148) noted that ‘it is something of a paradox that speech act theory emphasizes the uses of language, and in fact applies to utterances not sentences, but has depended largely on introspective judgments of isolated sentences’. Geis (1995) pointed out that acts such as requesting or inviting often develop over several interaction turns.

2000) probabilistic approaches, on the other hand, give an increasingly accurate surface description of empirical dialogues as successions of normalized ‘moves’ or ‘dialogue acts’ (Carletta et al., 1997; Stolcke et al., 2000); but the normalization of sequences as distinct ‘utterances’ also encourages an atomistic, ‘one segment, one act’ vision.<sup>6</sup>

Yet, as far as semantic-pragmatic representation is concerned, it is artificial and problematic to imagine that a request corresponds to a ‘block’ of signifier (§2.1) and to a ‘block’ of meaning (§2.2).

### 2.1 ‘Requests’—from signs to meaning

Spoken corpora show that requests are rarely composed of one clause or one simple clause-complex (though dialogue management requests might tend to correspond to monosegmental clauses or fragments). The majority of requests take the form of several increments of various syntactic, semantic and pragmatic types, often with repetitions of increments, interruptions from the co-interactants (and from the speakers themselves), embedded phases of negotiation, etc. The safest way to approach the problem is therefore to consider that *a priori* every request instantiation is likely to have a *discontinuous signifier* and *extensible boundaries*.

#### Requesting: a real-time process

Even assigning the beginning and the end of a ‘request’ in a linear transcription can prove difficult, as shown by some seemingly simple, supposedly straightforward ‘imperative’ requests:

- (4) Put a light in there // Put a light in there //  
(FUG.0h32m59s)
- (5) Put that gun down // Put that gun down // Now //  
(FUG.0h36m40s)

The police officer who utters (4) points successively at two different locations in a tunnel in which he is walking with his staff. It is therefore not a problem to say that the two clauses in this sequence correspond to two different requests. When some time later the police officer corners a fugitive and shouts (5), this ‘bjective’ analysis does not hold anymore: the officer does not want *two* separate actions of ‘putting the gun down’.

To consider the segment ‘Put that gun down’ as ‘a request’ would force us to consider the second

<sup>6</sup>That these ‘utterances’ may contribute to conversational or dialogue ‘games’ (Carletta et al., 1997; Levin et al., 1998, for instance) tends, in practice, to reinforce their atomistic character, despite the fundamental remark by Traum and Hinkelman (1992) on the divisibility of ‘utterances’.

segment, as well as ‘now’, as so many ‘requests’, since each of the three segments, in the situation, is pronounced in order to trigger an action. It is preferable to say that the three increments instantiate *one* same request. The reason is, that all throughout the speech sequence, the police officer has one result in mind only, and the fulfillment of his request at any moment of the sequence would in fact render the subsequent increments useless and even incoherent (which is not the case in (4)).<sup>7</sup>

The situation of a ‘surjective’ relationship between the speech increments and the acts they perform occurs in cases like (5) because the instantiation of action-requests rarely fits the ideal scenario according to which one ‘block’ *stimulus* triggers one clear *reaction* (or not), as summarized in the left part of Figure 1. Spoken language being a real-time process, interactants habitually take into consideration the current state of the world, checking whether their expectations have been met (and probably in a scalar not a polar way), and deciding whether more *stimulation* (seen as process) is necessary. In other words, the perlocutionary (the set of effects of the utterance) may influence the locutionary in return, as long as the world differs from what the speakers would like it to be; and as long as the speakers do not recognize the product of their intention in the world, they have to choose between several options: repeat, rephrase, modify the extent of, or abandon their requests. To account for this, one must consider request-instantiation not as an end-product but as a process, not as ‘act’ but as action in progress. This is summarized in the right part of Figure 1.

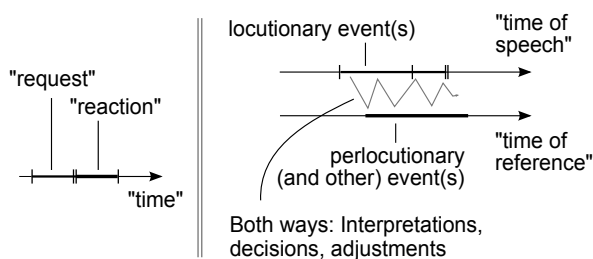


Figure 1: Requesting, a real-time process

<sup>7</sup>Still, it is common for a speaker to add request-related increments of specific types (e.g. stating motives) even after the addressee has started complying. This is because the speaker’s goal is usually not just to have the addressee fulfill the request: preserving or attaining a specific kind of relationship with one’s interactants (e.g. by sharing one’s reasons with them) is an objective in itself, which partly explains the ‘variability’ in request strategies (Bloomfield, 1933, §2.6) and sometimes even justifies the withdrawal of a request.

## The speech/act fracture

The problem of the instantiation of requests reaches in fact deeper than the ‘mere’ real-time calculation of an intention/effects (and /cost) ratio on the part of the speaker. Speech being linear, any request that uses spoken language for its instantiation will extend over a certain span. Often, the length of that span is primarily accounted for by the internal *complexity* of the request instantiation, even before it becomes relevant to assess the perlocutionary. That a radical fracture must be acknowledged between the speech elements used and the act performed is exemplified by (6), uttered in an emergency ward by a chief-doctor who is examining a patient, to a fellow doctor who is taking care of another patient a few meters away.

(6) Al / get over here / I need you // (FUG\_1h06m42s)

It is tempting to analyze (6) as a request preceded by a specification of its target and followed by an explanation of its motive. The problem with this ‘narrow’ analysis is that it forces major descriptive changes when possible variants are taken into consideration, such as (6′) and (6′′) (assuming that, in the situation, they could have produced the same effect). It seems indeed difficult to describe the increment ‘I need you’ as a ‘request’ in (6′) but as a mere explanation, ‘banalized’ by the presence of the imperative increment, in (6); and it seems equally difficult to hold that the vocative increment ‘Al’ falls within the scope of the request in (6′′) but outside of it in (6).

(6′) Al / I need you //

(6′′) Al //

The ‘narrow’ analysis presented above results from a vision of request instantiations in EITHER/OR terms within a limited range of clear-cut strategies (sometimes reduced to clause-types). However, the comparison of common requests such as (6) with their possible variants shows that the very idea of pinpointing one increment (usually a sentence) as *the* support of the request is taken at fault—so much so that the distinction between ‘direct’ and ‘indirect’ speech acts falls. Imperative clauses may retain a specificity compared with other segments (see further on), but the facts remain that (a) countless requests are instantiated by *several* increments, (b) a great number of these increments might suffice to instantiate the corresponding request alone, (c) none of these increments can claim to be *the* segment of speech that

performs the act', and therefore (d) all of the increments, far from excluding one another from the request instantiation, contribute to it.

From this perspective, the semantic-pragmatic segmentation of requests in spoken language consists in, first, identifying the increments involved in 'carrying out' the request and, second, determining the contribution of each increment to the whole. It seems doubtful, however, that the meaning of a request should be the simple, compositional addition of the meanings of its increments. One reason is that the increments of a request are often semantically and pragmatically heterogeneous (the three increments in (6) are by no means equatable), and their modes of contribution are therefore several. If heterogeneity is not a problem for the interactants, it is probably because they rely on a broader and more flexible vision of a request's *meaning* than is often acknowledged.

## 2.2 Requesting—from intention to signs

Why do speakers formulate requests? In most cases, the goal is not to actually witness an action. The speaker of (4) is not interested in 'seeing' his staff set up spotlights because he said so, nor is the speaker of (6) interested in 'seeing' 'Al' come over to her (if it were the case, motive increments such as 'I need you' would be incoherent, except maybe as ploys). The officer wants the tunnel to be lit, and the doctor wants immediate help from a colleague nearby. Each of these requests is therefore meant to bring about a specific *situation*. To be sure, the ultimate object of most requests is a desired state of the world—not an action; and the possible actions of the co-interactant(s) often have value not of themselves but primarily as a step towards the advent of that state.<sup>8</sup> A useful way to represent the speaker's mental context preceding a request instantiation is therefore to distinguish several 'worlds' roughly seen as static—the current world and the possible worlds (including the target world but also undesired worlds, maybe others still)—separated by a dynamic 'transition situation' which includes the possible actions of the co-interactants and the possibility for the speaker to stimulate them into enacting them.

Stating the reason that makes the target-world desirable ('I need you') or naming, with an imperative, the action that can bring about this world

<sup>8</sup>The idea of 'plan' below is close to that found in Perrault and Allen (1980, §3.3), but its treatment, from the perspective of a linguistic 'geo-strategy', will be different.

('Come here'), are two ways, *not mutually exclusive*, to reach the desired situation. Other increments could be added without 'exhausting' the meaning of the request, i.e. without instantiating a new or different one. Thus, the meaning of a request is often alluded to jointly by several elements which emphasize various parameters of the worlds or of the transition situation considered, along with elements referring to the process of requesting itself. Even when a speaker resorts to the imperative in order to call for an action, it is habitually clear to all that this action is to be enacted *in the name of* something else. The functioning of a request is therefore always *metonymic* (i.e. based on a radical speech/act fracture) in the sense that increments focus on one element or another and yet instantiate the whole request.

Requests might therefore be best described not as 'attempts... by the speaker to get the hearer to do something' (Searle, 1979, p. 13) but rather as *attempts to involve the addressee into a plan devised to reach a target world*—a plan defined especially (but never exclusively) by the projected action(s) of the addressee. In this perspective, the increments not naming the projected action are not seen as the conditions meant to help the addressee decode that action (in traditional pragmatic examples, these elements strangely disappear in the presence of the imperative); they are understood here as part of the general strategy of *sharing of a plan* that takes place when a request is instantiated (a 'sharing' which can, of course, be minimal).

## 3 Towards a 'constituency' of request instantiations

This section details how the reflections developed in §2 can be rendered operational in order to segment request instantiations. The aim here is to sort out the semantic-pragmatic 'constituency' of request instantiations, i.e. the system whereby higher (and often larger) units include lower (often smaller) ones. The task is therefore to find out which ranks are relevant for the study of requests and what primary distinctions should be acknowledged between the units of these ranks.<sup>9</sup>

<sup>9</sup>The reflection below is presented in a progressive way rather than as a complete nomenclature, both to show the empirical necessity of the distinctions acknowledged, and as a reminder that this model has not reached a definitive phase. Parallels may be found between the Universe, Manners and Phases described below and (respectively) the attentional, linguistic and intentional structures of Grosz and Sidner (1986), though with differences not discussed here.

### 3.1 Semantic distinctions

Part of what defines a request instantiation is, purely and simply, its ‘Universe’, understood as *what entities, concepts, qualities and relations are verbalized* (and regardless of the way they are). One level of description (at least) should therefore be dedicated to distinctions between the types of elements that appear in speech. Since the ‘Universe’ of speakers is not objective but rather peopled and structured by what makes sense in their world-views, some elements are given more prominence than others among the ‘worlds’ and the ‘transition situation’ assumed above.

In many request instantiations, the speaker refers to what could be called the ‘*head-action*’ of the request, i.e. the action that should lead most directly to the ‘target world’: ‘(Can you) tell me who...’ (2), ‘Put that gun down’ (5), etc. Regularly, however, actions are named which are not the *head-action* of the request. What is verbalized often corresponds either to a ‘*sub-action*’ of the head-action hoped for, or to a ‘*germane action*’ which is not a necessary subpart of the head-action but is, in the context, related to it and meant to trigger it. ‘Call’ names a *sub-action* in (7), as it is a necessary step within the head-action ‘tell’. ‘Look at it’ names a *germane action* in (8): with this utterance a fisherman asks a shark specialist to reconsider his judgment that the shark under examination (caught by the fisherman) is too small to qualify as the man-eater everyone is hunting for. Though ‘Looking’ is not properly a sub-action of ‘changing one’s judgment’ (or of ‘reconsidering’ it), it is supposed, in this context, to lead to that.

(7) ... You call Judge Rubin / you tell him I want a whole bunch of phone-taps... // (FUG\_0h29m37s)

(8) What / this is a big mouth / look at it // (JAW\_0h33m09s)

Many other verb-referents can be found which are not to be enacted by the addressee, such as ‘I was wondering’ in (2) or ‘You know’ in (1). The former is a (mental) process attributed to the first person (P1), the latter is a (mental) state attributed to the second person (P2). An additional difference is that ‘I was wondering’ is an assertion (of the occurrence of a reflection process) whereas ‘You know’, in this utterance, hesitates between the question and the assertion (as it often does). This difference relates to a second level of description: actions (and other elements) can be verbalized in different ‘*manners*’, for instance a *direct-*

*tive* or a *descriptive* manner. Following a traditional distinction, the former tells the addressee *to* do the action (through imperatives, performatives and maybe nominals, as in the army’s ‘Ateention!’) whereas the latter talks *of* or *about* the action (through various types of questions, assertions, exclamations, hypotheses, etc.). In (1), ‘It would probably be easiest...’ is the (descriptive) assertion of a judgment on ‘...if I just squeezed back there and poked around myself’, in which a sub-action and a higher-action (attributed to the first person, and to be *allowed*, not enacted by the second person) are verbalized as the (descriptive) evocation of a possibility (see Tables 2 and 3).

Other elements than actions are found in the verbalized ‘Universe’ of a request instantiation. People and objects are often named, some more than others. *Addressees* are crucial interactants as they are often hoped to become the agents of the projected head-actions, and they are therefore often named in separate increments (especially to attract their attention, as in (2), or to modalize the request). In a similar way, the notion of *head-object* can be useful to refer to those objects that occupy a central position in the representation of the head-action. Indeed, head-objects are so important on the ‘mind map’ of the speaker that they are commonly named without the action itself (‘Scalpel’ in an operating room, ‘The door!’, etc.). Sometimes, these objects are accompanied by other elements which help specify what is to be done, especially the *location* where the action is to take place or end. When people, objects, locations or other elements are verbalized outside the net of relationships found in clauses, they are often pointed at through speech, and the manner can be said to be ‘*indexical*’. Table 1 illustrates the concepts of ‘head’ action, object and location (FUG\_0h36m49s).

(9)	<i>Hands up</i>	<i>Over your head</i>	<i>Turn around</i>
Univ.	h-obj. <sub>R1</sub>	head-location(s) <sub>R1</sub>	h-action <sub>R2</sub>
Manner	idx.	idx.	directive
	REQUEST 1		REQUEST 2

Table 1: ‘Head’ actions, objects and locations

If the target-situation is desired (and if other situations, including the current one, are unwanted), it is usually because a change would be beneficial to someone or something (the speaker, the addressee, other people, institutions, moral principles, etc.). This explains why, quite often, *values* concerning the request plan are asserted (e.g. ‘It

would probably be easiest...’) or discussed (e.g. ‘Would that be alright with you?’). Elements of other types might be acknowledged in the Universe of the speaker, referring not to actions, participants, circumstances or values but to specific *meaning-contents* that can be given prominence when isolated as an increment. Thus, ‘Just’, in (2), does not refer to an entity but brings in the meaning of ‘restriction’—here a restriction bearing on the scope of the ultimate goal, in such a way that the request itself is attenuated. Often, the restriction bears directly on the head-action, as in (3).

### 3.2 Structural distinctions

One issue raised by the last remark is that of dealing with units of different ranks, as boundaries do not always coincide. ‘Just / for the week-end //’ is related to a verb in a preceding increment (‘getting an advance on my paycheck’) yet it appears after a *final* pause, as an afterthought. An approach favoring syntax might try to emphasize the relationship with the verb or the clause. A pragmatic alternative (or addition) can be proposed, underlining the fact that speakers (consciously or not) isolate some increments and join others. ‘Just’ carries the meaning of ‘restriction’; giving it prominence through prosody (by separating it both from what precedes and follows it) might therefore be an effective way of increasing the chances of a request to be fulfilled. Indeed, ‘Just / for the week-end //’ is not verbalized so much for the informational, referential specification it provides concerning the preceding verb, as for the way it restricts the ‘cost’ of fulfilling the *future* request (of asking for an advance on the salary) and therefore the *present* one (‘... tell me who I speak to...’).<sup>10</sup> The prosodic boundaries both signal and enact a reorganization of the roles and importance of verbalized elements under pragmatic considerations—a reorganization for which syntactic distinctions fail to account, and which might be erased or downplayed when increments are normalized into ‘utterances’.<sup>11</sup>

Still, ‘Just’ does not function alone. An accurate semantic-pragmatic description should be able to render the facts (a) that signs have meaning in isolation, (b) that they enter in meaningful larger syn-

<sup>10</sup>As ‘Mr Masry’ is the director of the firm, he can be expected to have a say in salary matters, which might influence his reaction to a request to name the office manager.

<sup>11</sup>In other words, this paper believes that functions are fulfilled a bit *below* (with ‘phases’, see §3.3) and a bit *above* (with ‘speech acts’, roughly Discourse Units as in Traum and Hinkelman (1992), and see §2) the level of ‘utterances’.

tactic structures (on which traditional corpus segmentation has focused) and (c) that prosodic cues often cut through these structures or fuse several of them,<sup>12</sup> an operation of (re)organization which is of semantic-pragmatic relevance.

Several types of meaningful units are therefore available.<sup>13</sup> They are treated here on three separate ranks (some of which might require subdivision to cover the whole structural complexity). The semantic ranks describing the verbalized Universe and the verbalizing Manner will commonly deal with whole increments as well as ‘sub-increments’ (e.g. ‘It would probably be easiest’ in (1)). The pragmatic ranks (see §3.3), on the other hand, will typically deal with whole increments, as what the speaker *does* through speech seems to be carried out by ‘phases’ which often fit into increment boundaries (or run over several increments). The general correspondence of phase boundaries with those of increments is strengthened by the observation that when several functions are fulfilled within the limits of one increment, they are usually fulfilled in a *syncretic*, not a successive fashion (though ‘phases’ may sometimes run on parts of increments only). Focusing on ‘functions’ of increments leads us to pragmatic distinctions.

### 3.3 Pragmatic distinctions

The increments uttered when instantiating a request are not just semantically and structurally heterogeneous (some assert judgments, others pinpoint objects or circumstances, etc.), they are also pragmatically heterogeneous: different types of ‘phases’ fulfilling different functions can usually be distinguished within a request instantiation.<sup>14</sup> Taking each increment one after the other, we can ask: what is the speaker trying to achieve with this increment with regard to the general request under way? do neighboring increments fulfill the same function? are several functions fulfilled by the same increment? if so, can the increment be divided into sub-increments corresponding to different phases, or are all the functions fulfilled syn-

<sup>12</sup>In ‘Do not smoke in here thank you very much //’ (JAW\_0h30m49s), fusion of an action-specifying phase with a ‘second answer’ (which normally follows a *positive* ‘first’ answer such a ‘Ok’) expresses the refusal of an alternative.

<sup>13</sup>None of these units need be ‘grammatical’ in the traditional, syntactic sense, as many types of *fragments* are in fact accepted in spoken language (Goldman-Eisler, 1968).

<sup>14</sup>The labels (e.g. *angling*<*calling*<*urging*) are meant to be ‘intuitive’. Their pragmatic relevance *vis-à-vis* the formal cues (word order, intonation, etc.) retained to describe speech elements with them is, of course, only assumed for English.

UNIVERSE	(major/head/high/sub/germane/next/ultimate) (mental/physical) ACTION/PROCESS/STATE (P{1-6}); TIME; PLACE; VALUE; 'RESTRICTION'; ... Notes: Modifiers before action, process or state specify their place on the mental world-map and their type. Modifiers afterwards attribute them to person. At this stage, values are not attributed to person.
MANNER	<u>EGOPHORIC</u> (sends back to speaker); <u>INDEXICAL</u> [INTERPELLATION/POINTING]; <u>DESCRIPTIVE</u> [ASSERTION/QUESTION/EVOCATION/RANGE/(...?) ( <i>of/on</i> ) ACTION/STATE/POSSIBILITY/CIRCUMSTANCE/JUDGMENT/REFLECTION/FEELING/...]; <u>DIRECTIVE</u> [IMPERATIVE/PERFORMATIVE/NOMINAL]; <u>OPERATIVE</u> (performs an action by itself, e.g. 'just', 'please', etc.); (...?) Notes: This level assesses the <i>semantico-structural</i> contribution; labels can be modified by 'Ambiguity', which is sometimes part of the speaker's strategy (see Table 3, (1)), and by 'Negative' ( <i>ibid.</i> , (3)).
PHASES	SPECIFYING/RESTRICTING/QUESTIONING ( <i>of</i> ) MOTIVE/ACTION/GOAL/SCOPE; ANGLING/CALLING/URGING ( <i>for</i> ) ATTENTION/FOCUS/EMPATHY/ACTION; (soft) PHATIC/FOCALIZING/ATTENUATING/MODALIZING/INTENSIFYING/... Notes: This level assesses the <i>pragmatic</i> contribution of increments. Label changes are to be expected.
DEPENDENCY	ATTENTION REQUEST; CONFIRMATION REQUEST; METAPRAGMATIC REQUEST; (...?) Notes: Request dependencies are described in Table 3 only when relevant.

Table 2: Preliminary ontology of levels of analysis and their labels (see Table 3 for application)

thetically by the whole increment? if the latter, are the functions distinct (a case represented by sign '&' in Table 3) or is one derived from another (in which case the position on a lower line without '&' represents derivation from functions on a higher line)?

Phases, through their 'vertical' relations with formal manners and their 'horizontal' interrelations, are useful to evaluate request strategies. One important difference between the directive and descriptive manners, for instance, is that in addition to specifying an action, directive manners conventionally convey an urge to enact it. Descriptive manners, as for them, are regularly accompanied by elements fulfilling other functions such as stating the value of the action (1) or questioning its possibility (2). Fine distinctions should also allow to compare, for instance, increments subtly 'angling' for attention (such as the throat clearing in (1)) and others more clearly 'calling' for it (the vocative in (2)), not to forget the cases where obtaining the addressee's attention is presented as a request in itself ('Now you listen' in (3)).

With this last remark, we are hitting upon an important pragmatic distinction: not all increments in a request participate in its instantiation equally. This is not just because different types of 'phases' must be acknowledged but also because, in some cases, these phases actually contribute to the request *via* their participation to the instantiation of a 'satellite', or 'dependent' request. True, requests for attention and confirmation are commonly found as 'independent' requests (for in-

stance in a classroom), but they often serve the purpose of ensuring the felicity of a 'main' request, as in (3). 'Now you listen' and 'You got that?' are requests in their own right,<sup>15</sup> nevertheless, these requests would have no *raison d'être* without the main request not to follow the speaker. A layer can therefore be added in the tables to account for request 'Dependency'; and the head-action of the *main* request gains a new status, as '*major*' action in the request plan.<sup>16</sup>

#### 4 Limitations and prospects

The model presented here has not yet reached a state of maturity where its reliability as an annotation scheme can be tested. Fine-tuning of the distinctions, and clear decision-trees for each rank, are among the next necessary steps. One theoretical limitation is that this approach, in its current form, does not cover the use of metaphoric language and more generally the *lexical* contribution of a number of elements (for instance, the non-professional and vague verb 'to poke around' in (1) is not chosen by chance instead of, say, 'to search for the legal records my firm needs'). As important is the need to take prosody into fuller

<sup>15</sup>The co-speaker's 'Yeah' following the latter is not only an ANSWER but also an AGREEMENT/ACCEPT (or COMMIT), in terms of the SWBD-DAMSL taxonomy (Jurafsky et al., 1997; Stolcke et al., 2000). The general duality affecting 'check questions' was noted in Core and Allen (1997).

<sup>16</sup>Clarification is needed of the Phase/Dependency boundary, i.e. of the criteria used to decide when 'phases' of a request acquire the status of 'dependent request' (of which some uses of 'Come on' and 'Do it' illustrate another type, that of a 'metapragmatic' request to fulfill the main request).

(1)	<i>Err hmm</i>	<i>You know</i>	<i>It would probably be easiest</i>	<i>if I just squeezed back there and poked around myself</i>	<i>Would that be alright with you?</i>
Universe	-	state (P2)	value	sub-act° + high-act° (P1)	value
Manner	egophoric	descriptive Ambig. Ass°/Q°	descriptive Assert° Judgment	descriptive Evocat° Possibility	descriptive Quest° Judgment
Phases	att° angling	empathy angling	specifying motive & specifying goal		assent angling
REQUEST (to let speaker go look for files herself)					

(2)	<i>Mr Masry?</i>	<i>(-Yeah)</i>	<i>I was wondering</i>	<i>can you</i>	<i>tell me who I talk to</i>	<i>about maybe getting an advance on my paycheck</i>	<i>Just</i>	<i>for the week-end</i>		
U	addressee		mental proc. (P1)	state (P2)	head-a°	next a° (P1)	ultimate a° (P1)	'restr°'	moment	
M	indexical Interpellat°		descriptive Assert° Reflect°	descr. Q° Poss.	descr. Evoc° P.	descr. Evoc° P.	descriptive Evocat° Possib.	operative Restrict°	descr. Evoc° Circ.	
Ph	att° calling		soft focalizing & soft modalizing	specifying goal=motive & head-act° specifying & Calling			restricting scope attenuating modalizing			
D	ATTENT° R.	✓	REQUEST (to tell speaker how (and if) possible to get advance on her paycheck)							

(3)	<i>Now</i>	<i>you listen</i>	<i>I don't give a damn</i>	<i>which way you go</i>	<i>Just</i>	<i>don't follow me</i>	<i>You got that?</i>
Univ.	time	head-act° <sub>R'</sub>	value (feeling)	act° (P2)	'restr°'	major action	sub-action <sub>R''</sub>
Man.	idx.	directive Imp.	descriptive Assertion Feeling	descriptive Range Act° <sub>s</sub>	operative Restrict°	directive Imp. (neg.)	descriptive Quest° Act°
Ph.	focus angl.	& spec. act° & urging	Expressing concession modalizing phase		:	specifying act° & urging	spec. s-act° & questioning
Dep.	ATTENTION R.		REQUEST FOR ACTION (not to follow speaker)				CONFIRM° R.

Table 3: Description of examples (1), (2) and (3)

account and to include nonverbal cues. Another issue is the fact that repair, backchannel and overlapping tend to be more common in spontaneous speech than in films (work in preparation); these phenomena (all of which can be of pragmatic significance in the context of request-formulation), as well as turn-taking, must be better integrated.

On the plus side, this approach has the advantage of trying to bridge the gap, with strong *empirical* emphasis, between 'emic' parameters such as the speakers' beliefs, desires and intentions, and 'etic' cues from the signifier (Pike, 1954; Blum-Kulka, 1981; Reiss, 1985; Jurafsky, 2004). By focusing on the contribution of increments to the construction of meaning, and by running statistics to reveal which types of increments are used by speakers in which context and according to which concatenation patterns, we should eventually be able to draw a picture of the 'strategies'—conscious or routinized—used when requesting.

## 5 Conclusion

Traditional speech act theory rests primarily on the structure of isolated sentences. However, at least as far as requests are concerned, speakers tend to express themselves with several increments, heterogeneous both in nature and function. 'Cue-based' approaches designed to *recognize* atomic acts can give accurate descriptions of the speech surface; but the treatment of each unit as 'act' tends to blur the deeper interrelations. The approach presented here, based on the loose inclusion of a lower semantic level under a higher pragmatic level via a structural level, suggests that, as far as the representation of spontaneous spoken language is concerned, gains might be made by broadening the scope of dialogue acts and 'lowering' the aim from the identification of distinct acts to that of the means of their instantiation.

## Acknowledgments

I wish to thank Pierre Cotte, Frank Alvarez-Pereyre, Ruth Shalev and an anonymous reviewer for useful comments and discussions.



## References

- John L. Austin. 1962. *How to do things with words* [edited by J. O. Urmson]. Clarendon Press.
- Leonard Bloomfield. 1933. *Language*. Holt & Co.
- Shoshana Blum-Kulka. 1981. The study of translation in view of new developments in discourse analysis: The problem of indirect speech acts. *Poetics Today*, 2(4):89–95.
- [BRO] *Erin Brockovich*. 2000. Directed by Steven Soderbergh. Written by Susannah Grant. Jersey Films.
- Jean Carletta, Amy Isard, Stephen Isard, Jacqueline C. Kowtko, Gwyneth Doherty-Sneddon, Anne H. Anderson. 1997. The Reliability of A Dialogue Structure Coding Scheme. *Computational Linguistics*, 23(1):13–31.
- Mark Core and James Allen. 1997. Coding dialogs with the DAMSL annotation scheme. In *Working Notes of the AAAI Fall Symposium on Communicative Action in Humans and Machines*, pages 28–35.
- William A. Corsaro. 1977. The clarification request as a feature of adult interactive styles with young children. *Language in Society*, 6:183–207.
- [FUG] *The Fugitive*. 1993. Directed by Andrew Davis. Written by Roy Huggins, David Twohy and Jeb Stuart. Warner Bros. Pictures.
- Michael L. Geis. 1995. *Speech acts and conversational interaction*. Cambridge University Press.
- Frieda Goldman-Eisler. 1968. *Psycholinguistics: Experiments in spontaneous speech*. Academic Press.
- H. Paul Grice. 1975. Logic and conversation. In P. Cole and J. Morgan, editors, *Syntax and semantics 3: Speech acts*. Academic Press.
- Barbara J. Grosz and Candace L. Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204.
- [JAW] *Jaws*. 1975. Directed by Steven Spielberg. Written by Peter Benchley (based on his novel) and Carl Gottlieb. Zanuck/Brown Productions – Universal Pictures.
- Daniel Jurafsky, Elizabeth Shriberg and Debra Biasca. 1997. *Switchboard SWBD-DAMSL labeling project coder's manual, Draft 13*. TR 97-02, University of Colorado Institute of Cognitive Science.
- Daniel Jurafsky and James H. Martin. 2000. *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Prentice Hall.
- Daniel Jurafsky. 2004. Pragmatics and computational linguistics. In L. R. Horn and G. L. Ward, editors, *The handbook of pragmatics*, chapter 26. Blackwell.
- Alessandro Lenci. 1994. A relevance-based approach to speech acts. In E. Fava, editor, *Speech acts and linguistic research*. Proceedings of the workshop, July 15–17, 1994, Center of Cognitive Science of New York at Buffalo. Nemo, 1995.
- Lori Levin, Ann Thymé-Gobbel, Alon Lavie, Klaus Ries and Klaus Zechner. 1998. A discourse coding scheme for conversational Spanish. In *Proceedings of ICSLP 98*, 6:2335–2338.
- C. Raymond Perrault and James F. Allen. 1980. A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics*, 6:167–182.
- Kenneth L. Pike. 1945. *The intonation of American English*. University of Michigan Publications. Reprinted in part as Chapter 3 in D. Bolinger, editor, *Intonation*, Penguin Books, 1972.
- Kenneth L. Pike. 1954. *Language in relation to a unified theory of the structure of human behavior*. Summer institute of linguistics.
- Matthew Purver, Jonathan Ginzburg and Patrick Healey. 2003. On the means for clarification in dialogue. In J. C. J. Van Kuppevelt and R. W. Smith, editors, *Current and New Directions in Discourse & Dialogue*. Kluwer Academic Publishers.
- Matthew Purver. 2004. *The theory and use of clarification requests in dialogue*. Ph.D. thesis. King's College, University of London.
- Nira Reiss. 2004. *Speech act taxonomy as a tool for ethnographic description* (Pragmatics and Beyond VI:7). John Benjamins.
- Kepa J. Rodríguez and David Schlangen. 2004. Form, intonation and function of clarification requests in German task-oriented spoken dialogues. In *Proceedings of Catalog '04* (The 8th Workshop on the Semantics and Pragmatics of Dialogue, SemDial04), pages 101–108.
- John R. Searle. 1969. *Speech acts: An essay in the philosophy of language*. Cambridge University Press.
- John R. Searle. 1979. *Expression and meaning: Studies in the theory of speech acts*. Cambridge University Press.
- Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Marie Meteer and Carol Van Ess-Dykema. 2000. Dialog act modeling for automatic tagging and recognition of conversational speech. *Computational linguistics*, 26:98–105.
- Michael Stubbs. 1983. *Discourse analysis: The sociolinguistic analysis of natural language* (Language in Society 4). Basil Blackwell Ltd.
- David R. Traum and Elizabeth A. Hinkelman. 1992. Conversation acts in task-oriented spoken dialogue. *Computational Intelligence*, 8(3):575–599.