Functional constraints on sentence processing: A cross-linguistic study

ELIZABETH BATES*
University of Colorado

SANDRA McNEW
University of Colorado

BRIAN MacWHINNEY
University of Denver

ANTONELLA DEVESCOVI
University of Trieste

STAN SMITH
University of Colorado

Abstract

English and Italian provide some interesting contrasts that are relevant to a controversial problem in psycholinguistics: the boundary between grammatical and extra-grammatical knowledge in sentence processing. Although both are SVO word order languages without case inflections to indicate basic grammatical relations, Italian permits far more variation in word order for pragmatic purposes. Hence Italians must rely more than English listeners on factors other than word order. In this experiment, Italian and English adults were asked to interpret 81 simple sentences varying word order, animacy contrasts between the two nouns, topicalization and contrastive stress. Italians relied primarily on semantic strategies while English listeners relied on word order—including a tendency to interpret the second noun as subject in non-canonical word orders (corresponding to word order variations in informal English production). Italians also made greater use of

1: This research was carried out with the cooperation of the Institute of Psychology, CNR, Rome, Italy, with further funds provided by National Science Foundation Grant # BNS 7905755 to MacWhinney and Bates.

*Reprint requests should be sent to Elizabeth Bates, Psychology Department, Muenzinger Building, University of Colorado, Boulder, Colorado 80309, U.S.A.
topic and stress information. Finally, Italians were much slower and less consistent in the application of word order strategies even for reversible NVN sentences where there was no conflict between order and semantics. This suggests that Italian is 'less' of an SVO language than English. Semantic strategies apparently stand at the 'core' of Italian to the same extent that word order stands at the 'core' of English. It is suggested that these results pose problems for claims about a 'universal' separation between semantics and syntax, and for theories that postulate a 'universal' priority of one type of information over another. Results are discussed in the light of the competition model, a functionalist approach to grammar that accounts in a principled way for probabilistic outcomes and differential 'weights' among competing and converging sources of information in sentence processing.

Italian and English provide an interesting contrast with regard to a controversial issue in psycholinguistic research: the boundary between grammatical and extra-grammatical knowledge in sentence processing. Although these two languages are formally quite similar (i.e., both are Subject-Verb-Object word order languages without case inflections to indicate basic grammatical relations), they are informally very different. In particular, Italian allows considerably more variation in word order to express topic/focus information; permissible word order variations are constrained in large measure by disambiguating information in the context.

We assume that listeners are able to bring all their cognitive resources to bear on the problem of sentence interpretation (Clark, 1978). The question is, how much of this knowledge must be described within the grammar? From the point of view of transformational grammar, Fodor, Bever and Garrett (1974, pp. 368–370) have suggested that language comprehension may short-circuit our knowledge of grammar altogether, functioning instead through some independent processing strategies that must be described in a separate theory of performance. When listeners apply semantic and perceptual strategies to interpret English sentences, they are exercising a competence-to-perform which is related only indirectly to their tacit knowledge of grammatical relations. Presumably, when Italians use contextual information to interpret pragmatic word order variations, these strategies would also belong not in the grammar but in a performance component. Forster (1979) has carried this analysis even further, suggesting that grammars evolved as mental entities in direct opposition to semantic strategies, to help us to say unlikely things when necessary (pp. 55–56):

'The whole point of a language having a syntax is to provide a clear and unmistakable indication of the correct interpretation of the sentence... Any move to allow
the syntactic processor to be influenced by pragmatic factors works against the fundamental purpose of syntax. In fact, one might surmise that the evolution of syntax has been influenced by the degree to which it successfully guards against errors introduced by a consideration of pragmatic and semantic facts. Hence, it has always seemed strange to me that anyone should expect to find evidence for semantic influences on syntactic processing. To be sure, one might expect to find evidence that syntactic indicators are often overlooked or overridden when the appropriate semantic interpretation of the sentence seems obvious from the context, or from a consideration of the lexical items used... But these are effects of pragmatic factors on message processing, not on syntactic processing.'

Gleitman and Gleitman (1970) take a slightly different approach to this boundary issue. They also place semantic strategies outside the grammar, as a resource that is called upon when syntactic strategies fail. However, they are more optimistic than Fodor et al., concerning the relevance of processing data to the study of grammatical knowledge. To deal with difficult boundary phenomena, they offer two distinctions:

1. Core versus penumbral rules in the grammar
2. Grammatical versus semi-grammatical processing.

The core/penumbra distinction grew out of Gleitman and Gleitman's own work on the interpretation of nominal compounds by naïve English listeners. Some of their subjects interpreted novel compounds 'The bird boot house' according to a systematic set of recognized English order and stress rules. Hence, they might paraphrase 'the bird boot house' as 'a boothouse that belongs to birds', while interpreting 'the bird boot house' as 'a house for birdboots'. However, other subjects failed to show any control of these rules, ignoring order and stress to interpret both versions with the same set of semantic strategies (e.g., 'a bird house with boots in it'). Gleitman and Gleitman suggest that nominal compound rules belong to the 'penumbra' of English:

'We will use the term 'core' to refer to central aspects of the grammatical system and 'penumbra' for the more peripheral features... We might conclude that penumbral rules are simply additional rules of the grammar, restricted though they are in range and frequency of use. Elite speakers 'have' them and others do not.' (p. 158)

If a speaker does not have a set of penumbral rules, or if those rules are unstable and difficult to apply, he still need not by-pass the grammar altogether. Instead, deviant forms may be interpreted with an ancillary device called a semi-grammar. This semi-grammar is a set of procedures for relating deviant utterances back to the grammar by calculating their degree of resemblance to well-formed structures. Chomsky (1961) has put forth some
related proposals concerning ‘degrees of grammaticality’ (see Moore, 1977, for further discussion). Given a sentence like ‘The flower picked the boy’, listeners have two choices: (1) they can ignore the semantic implausibility of the combination and assume that standard word order is intended, or (2) they can assume that there has been a violation of word order and that a semantically-plausible message was intended. In either case, there are specifiable operations for relating this utterance to a more standard English sentence (Note: in Chomsky, 1965, this sentence represents not just a semantic anomaly, but a violation of the syntactic selection-restriction that animate verbs must take animate subjects).

If basic word order is deformed even further, as in ‘The flower the boy picked’, the listener may still be able to apply operations in the semi-grammar to make an interpretation. For example, Bever (1970) has shown that children may use a ‘first noun strategy’ that leads them to interpret passive sentences as active ones. Forster (1979) has proposed that a similar first noun strategy might be used by adults:

‘It might be easier to ‘guess’ the meaning of doctor patient cure than the meaning of patient doctor cure; in the former case, the initial position of doctor may establish it as having the semantic function of actor.’ (p. 53)

Both of Gleitman and Gleitman’s distinctions underscore the relevance of processing data to theories of grammar. Reaction time and individual variation in response strategies may tell us about the core versus peripheral status of rules in the grammar. Interpretations of deviant forms may tell us how the grammar can be ‘stretched’ to account for phenomena just outside its boundaries. However, contextual effects and semantic strategies still remain outside the grammar, in a separate performance component that is a last resort for conscientious listeners.

In Italian, semantic strategies may play a central role even in the interpretation of very simple sentences. If we can demonstrate a ‘core’ use of semantic strategies even for elite Italian listeners, then serious questions are raised about the universality of models that place strict boundaries between semantics and the grammar.

In their theory of cognitive grammar, Lakoff and Thompson (1977) extend the boundaries of the grammar still further to include virtually all systematic aspects of sentence processing.

‘We believe that there is a direct and intimate relation between grammars and mechanisms for production and recognition. In fact, we suggest that GRAMMARS ARE JUST COLLECTIONS OF STRATEGIES FOR UNDERSTANDING AND PRODUCING SENTENCES. From this point of view, abstract grammars do not have any separate mental reality; they are just convenient fictions for representing certain processing strategies.’ (p. 295)
The Lakoff and Thompson approach falls within a class of theories that Bates and MacWhinney (1979, 1981) describe as functionalist grammars. In a functionalist theory, there is no competence other than competence-to-perform. Semantic, pragmatic and perceptual-mnemonic strategies must all be represented directly in the grammar, i.e., in descriptions of the elements, categories and operations that mediate between meanings and the particular surface forms of a given language. Contextual strategies for the interpretation of Italian variations in word order would be a central part of a functionalist grammar.

Given the diversity of theories and paucity of data of processing strategies, cross-linguistic studies are particularly valuable for at least narrowing down the range of potential ‘universals’ in sentence processing. We share the view of many psycholinguists that more progress can be made in studies of language processing if we are informed by the principles and assumptions of some kind of grammatical theory. For this reason, the present study was designed and interpreted within the framework of the competition model (Bates and MacWhinney, 1981; MacWhinney, 1981). This model is only one of many possible functionalist theories; furthermore, it is primarily a processing theory, which does not yet incorporate the descriptive detail required for a full competence model. However, as a performance grammar it provides breadth and flexibility in accounting for a variety of phenomena from different languages and different performance domains.

The present study was designed as a test of the explanatory adequacy of the competition model in accounting for the comprehension of simple sentences. Before describing the method and results of this study, we first need to consider (1) a description of differences that make an English-Italian comparison interesting, and (2) a brief characterization of the competition model and the kind of predictions it makes for the English-Italian contrast.

Some Facts about English and Italian

In traditional language taxonomies (e.g., Greenberg, 1961), Italian and English belong to the same basic category. That is, they are both non-agglutinative word order languages with a basic or unmarked order of SVO (subject-verb-object). Both languages do make case distinctions among personal pronouns (e.g., I versus me in English). Also, both languages mark the verb to agree with the subject in person and number. With regard to agreement, the Italian verb morphology system is richer and less ambiguous than the corresponding forms in English as follows.
Singular  |  I buy          | Io compro  
          | You buy (informal) | Tu compri 
          | You buy (formal)   | Lei compra  
          | S/he buys         | (Lei/lui) compra 

Plural    | We buy         | Noi compriamo 
          | You buy (informal) | Voi comprate 
          | You buy (formal)   | Loro comprano 
          | They buy         | Loro comprano 

English only marks third-person singular as distinct (*buys*), thus almost always relying on subject pronouns. In many cases, the corresponding Italian verbs are transparent in person without subject pronouns, having distinct inflections for each person-number form. However, ambiguity is possible in Italian as well, i.e., when there are two third-person references of the same number, or in sentences mixing third person and second-person formal reference. Thus, even though Italians can rely more often on disambiguating cues from verb morphology, morphological ambiguities are possible in both languages. Hence we might expect word order to be the primary vehicle for indicating basic grammatical relations in either case.

On the other hand, there are some formal and informal differences between the two languages in the way that word order is used. One of these involves variations in word order within relative clauses. To illustrate, both of the following sentences are acceptable in spoken or written Italian:

(1) Siamo andati in un posto dov’è mangiano i selvaggi.
    (We)-went to a place where (they)-eat the savages.

(2) Siamo andati in un posto dov’è i selvaggi mangiano.
    (we)-went to a place where the savages (they)-eat.

Sentence (1) is an utterance which Bates (1976) reports hearing at a cocktail party in Rome. In principle, both (1) and (2) are ambiguous in Italian between a reading in which ‘savages’ are doing the eating *versus* a reading in which ‘savages’ are being eaten. The second reading would be most unlikely in most contexts, but Bates reports that this was in fact the listener’s mistaken interpretation. The reason for this rare misfire is, according to Bates, because the statement was made at a gathering of anthropologists who had worked among formerly cannibalistic cultures. This example illustrates how contextual factors may be used in Italian to deal with the inherent ambiguity of alternative orders. In English, only one word order is possible in relative clauses. To provide the cannibalistic reading, we would have to use some kind of a passive constructions (e.g., ‘where the savages get eaten’).
In informal discourse, word order differences between English and Italian are much more extreme, extending to the order of basic constituents in simple active declarative sentences with no embeddings. Starting with English, we do occasionally find 'left-dislocated', OSV structures like the following:

(3) John I like, but Roger I can't stand.

But such constructions are restricted to informal settings (i.e., they are not appropriate in written texts). Also, English speakers often tend to double-mark such topicaized objects with pronouns in the main clause, e.g.,

(4) That guy I told you about at the office, I can't stand him.

In this case, (4) is really not a violation of word order, i.e., an OSV construction. A better characterization would be O,SVO with a perceptible pause between the O fragment and the full clause.

Another type of reordering that does occur in English is the right-dislocation such as the following:

(5) Likes eggcreams, John
(6) Really gets on my nerves, old Roger.

We would call these VOS constructions. However, another characterization might be that the subject of the main clause was deleted before the verb and then added on to the end of the clause as an afterthought: (S)VO,S. There may be a pause between the verb and the afterthought, and the right-dislocated subject typically receives very minimal stress. If we accept this characterization, then (5) and (6) do not have to be considered as VOS exceptions to the basic English SVO order. At any rate, like left-dislocations, these right-dislocations also are possible only within informal discourse.

It is interesting that in English pragmatic reorderings we find OSV (left-dislocation) and VOS (right-dislocation) but not SOV and VSO. Take the following examples:

*(7) I John like, but I Roger can't stand.
* (8) Likes John eggcreams.

Hence there seem to be structural constraints on these informal distortions of basic SVO order in English. SOV structures occasionally occur in poetic forms, e.g.,

(9) In Xanadu did Kubla Khan a stately pleasure dome decree

But they sound archaic and inappropriate in modern English conversation.
In Italian, on the other hand, all possible orders of subject, verb and object can and do occur in informal discourse, under certain conditions. Some examples can be seen in the following hypothetical but representative sequence from a restaurant conversation:

**OSV (10)**  
La pastasciutta Franco la prende sempre qui.  
Pasta Franco (it) orders always here.

**VSO (11)**  
Allora, mangio anche io la pastasciutte.  
Well then, am eating also I pasta.

**VOS (12)**  
Ha consigliato la lasagna qui Franco.  
Recommended the lasagna here Franco.

**OVs (13)**  
No, la lasagna l'ha consigliata Elizabeth.  
No, the lasagna (it) recommended Elizabeth.

**SOV (14)**  
Allora, io gli spaghetti prendo.  
In that case, I the spaghetti am having.

Although such sequences are quite common, they have so far eluded formal description. For one thing, it is difficult for Italians to judge such orders one at a time, out of context, since they are apparently conditioned by the pragmatics of connected discourse. One solution might be to exclude alternative orders from the grammar altogether, to be explained within some sort of semi-grammar for dealing with deformations and violations of basic rules. Evidence for excluding them from the grammar comes from the fact that they typically occur with abnormal stress patterns, and with clitic pronouns to double-mark those objects that appear out of canonical position (e.g., (10) and (13) above). Another factor to consider is that subject deletion is quite common in informal Italian. The subject of main verbs is left out of up to 70% of the possible cases (based on estimates from Bates, 1976). Since deleted subjects are quite grammatical, a sentence like (12) could be described at the surface level as the following:

(S)VO,S (15)  
(Lui) ha consigliato la lasagna qui, Franco.  
(He) recommended the lasagna here, Franco.

Taking this line of argument, alternative orders might be viewed as combinations of sentence fragments with deleted elements, rather than true SOV, VOS, VSO and OSV sentences.

On the other hand, some Italian grammarians have argued that all or most of the alternative orders are completely acceptable without special stress or clitics, under certain conditions of givenness and newness (Antinucci, 1977; Duranti and Ochs, 1979). They do not deny that the basic, unmarked order
of Italian is SVO, since that order can be used under a wider (though not unlimited) range of pragmatic conditions. But they do claim that alternative orders belong in the grammar, and that the grammar should in turn provide mechanisms for describing topicalization as a basic determiner of constituent order in surface structure.

The contrast between English and Italian in flexibility of word order raises two interesting questions for a study of sentence comprehension. First, regardless of the status of pragmatic reorderings in formal characterizations of the grammar, what effect do they have on language use? We know that these forms occur in spontaneous production. Do they also play a role in comprehension? The second question regards the relative contribution of pragmatic, semantic and syntactic factors in sentence interpretation. Italians apparently listen to massive distortions of SVO order all the time—a fact which, according to Bates (1976), actually delays the acquisition of basic word order in Italian children. To interpret these varying orders, listeners must rely on many different sources of information including lexical contrasts (e.g., that lasagnas don’t eat people), contrastive stress, and knowledge of previous discourse. For this reason, a performance grammar of Italian may look very different from a performance grammar of English. The two languages provide a strong test of would-be universal hypotheses about the role of syntactic factors in sentence comprehension.

The competition model

The following tenets from the competition model (Bates and MacWhinney, 1981) are relevant to the present study.

1. Competition

A first tenet is that functional constraints must compete for control of the limited resources of the acoustic-articulatory channel. Ideally, in the absence of resource limitations, natural languages might model the entire grammar after the one-to-one form/function mappings of the lexicon. For example, case inflections on nouns could be used exclusively to convey basic semantic relations, while word order variations are used only in the service of discourse functions like topicalization. This is what Bates and MacWhinney call a ‘divide the spoils’ system. However, given the four kinds of surface devices available in natural languages (e.g., lexical items, morphological markings on lexemes, word order and prosody), with concomitant memory and perceptual constraints among these interacting devices, we assume that surface
resources are insufficient to meet communicative purposes using nothing but a one-to-one system. Some 'doubling up' is necessary for efficient processing. Furthermore, some surface devices are inherently 'better' than others: more perceivable, more easily accessed, more rapidly processed. And some functional categories are 'more important' than others: higher frequency, higher information value. These disparities lead to further competitive pressures on the mapping apparatus (i.e., the grammar) that mediates the two levels.

2 Coalition

The second tenet is that natural languages exploit situations of natural overlap among functions, assigning certain surface devices not to one function but to a coalition of functions that engage in 'peaceful coexistence' much of the time. An example relevant to the present study regards the relationship between topic and agent (setting aside for the moment the fact that these two categories must also have complex internal structures, which may in turn involve another level of competition). Both of these categories are strong vectors (i.e., a determining force, with measurable magnitude and direction) in natural discourse, so that they can both claim control over central and salient surface devices in the grammar. However, it is also the case that topic and agent overlap much of the time in natural discourse (Li and Thompson, 1976; MacWhinney, 1977). Natural languages may recognize and institutionalize this high-probability overlap by creating surface devices that are governed by a joint category like topic-agent.

In addition, natural languages may also form coalitions of surface forms, in which forms that previously served separated functions come to operate together as a block. An example of such a device is the surface role of subject. As described by Keenan (1976), the role of subject is in fact a coalition of disparate surface devices that all tend to be assigned to the same underlying element (i.e., topic-agent). Membership in this surface coalition may vary from one language to another. For example, in both English and Italian the subject coalition includes preverbal, sentence-initial position in the unmarked order of simple active declarative sentences, plus agreement with the verb in person and number, and nominative case marking in the pronoun system. The complete grammar that results from such evolving alliances would be comprised of a large and interrelated set of surface categories containing coalitions of devices, and functional categories containing coalitions of meanings and intentions.
3. Category membership

By the above argument, the decision to assign a surface role to a particular element would be triggered by the category membership of that element at the functional level. For example, within a given intended meaning structure, that element which provides the 'best fit' to the topic-agent category would be assigned the surface role of subject. As summarized in Bates and Mac-Whinney (1981), a number of problems concerning categorization processes are solved if we assume that grammatical categories (at both the surface and functional levels) are structured in a manner similar to the category structures proposed in Rosch’s theory of prototypes (Rosch, 1977). Categories are defined in terms of an idealized ‘best member’, or prototype, which may have a heterogeneous internal structure (as with the coalitions described above). Category membership is assigned to new instances on the basis of degree of resemblance to the prototype, ranging from central ‘best instances’ to peripheral instances at the ‘fuzzy’ outer boundaries of the category. Applied to grammatical categories, this approach has advantages in explaining some syntactic differences between ‘good’ and ‘bad’ members of the same form class, e.g., stative versus active verbs (Ross, 1972). The potential utility of prototype theory in grammar has been discussed by Braine and Wells (1978), deVilliers (1981), Lakoff (1977), and Schlesinger (1977), among others.

4. Breakdown of coalitions

When the high-probability situation of functional overlap breaks down, languages must provide alternative (albeit less efficient or desirable) means for expressing the competing functions separately. For example, even though the surface role of sentence subject in active declarative sentences is best filled by a topic-agent, it may occasionally be the case that we want to topicalize the patient of the verb while at the same time identifying the agent. There are two basic kinds of solutions to such situations: (a) adopt an alternative surface device, or (b) break the original surface coalition down into its component parts and assign those parts separately. For example, in the patient-topic situation just described, the English alternative is passivization: a unified surface subject role is assigned to the topic (i.e., preverbal position, agreement with the verb, nominative case pronouns where relevant); but the structure of the verb is changed, and the agent is expressed through a by-clause. This alternative is available in Italian as well. However, the most frequent response to the patient-topic situation in Italian is to break the subject role into separate devices: first position is assigned to the topic, but agreement with the verb is assigned to the agent (resulting in the
word order variations described earlier). Another way to put this is that the unity of the surface subject coalition, i.e., its tendency to operate as a block, is stronger in English than it is in Italian.

5. Vector weighting

When a functional coalition breaks down, something has to determine which element in the coalition 'wins' access to particular surface privileges (e.g., which one is assigned first position). We assume that coalitions of functions operate like coalitions of parties in a parliamentary system: they maintain their separate identity and their constituencies, so that distribution of resources when an alliance is suspended is determined by relative power or strength in different sectors. In a performance grammar based on the competition model, individual functions within a coalition carry canonical weights or vectors that state their degree of association with a given surface form (including one or more components of a surface coalition). Analogous to the feature weightings proposed by Tversky (1977) to explain similarity judgments in visual perception, these weights may vary to some degree depending on the context. But they are relatively predictable across contexts. In languages with similar surface structures, the relative weights of underlying functions may vary. Hence in Italian first position may be more strongly associated with topic, while it is more strongly associated with agent in English. However, in both languages the 'best' subject and the strongest candidate for first position would be a prototypic topic-agent.

6. Conventionalization

What we have described so far is essentially a 'fuzzy' performance grammar, in which the line between probabilistic tendencies and determinate rules becomes a matter of degree.

Within generative grammars, the distinction between probabilistic and deterministic tendencies has usually been described with optional versus obligatory rules. Obligatory rules must apply in every context that meets their minimal structural specifications. Optional rules may apply in some contexts, but not in others. Cederberg and Sankoff (1974), and Labov (1969) have argued that optional rules should not be viewed as discrete, 'yes or no' statements, but as variable rules which specify the probability that a given operation will apply.

In the competition model, the optional/obligatory distinction is captured diachronically in terms of a process of conventionalization. This is an historical process in which natural languages increase (a) the vector weights that
hold for certain form/function mappings, and/or (b) the unity of certain surface form coalitions (e.g., surface subject), to the point where violations never (or rarely) occur in natural language use. Hence the probabilistic information that Cedergren and Sankoff treat as variable rules, would be represented here as the weight or strength of a given form/function mapping. At a given moment in the history of language (or in the acquisition of that language by a child), two apparently obligatory or conventionalized tendencies may actually have slightly different psychological weights (e.g., 100% versus 97%). We would not detect such a hypothetical difference under natural conditions. But under experimental conditions, in which two apparent rules are set into direct competition, a difference in relative weights may emerge.

We can illustrate such language-specific differences in degree of conventionalization with the so-called ‘animacy hierarchy’ or ‘closeness-to-ego’ principle (Cooper and Ross, 1975; Dick, 1978; Kuno, 1975). This is a probabilistic tendency across natural languages for word order to reflect a hierarchical ordering from most like the speaker (i.e., first and second versus third person), most human, most animate, and so on, to the most inactive and inanimate elements. Languages vary in the degree to which this universal tendency is conventionalized within the grammar. For example, English permits inanimate instruments to serve as subjects (e.g., ‘The knife sliced the salami’) while Dutch does not (Snow, 1979). And Navajo has conventionalized the entire hierarchy, including some power distinctions peculiar to Navajo culture, into an obligatory system of word order. Hence the ‘degree of conventionalization’ of the animacy hierarchy increases from English, to Dutch, to Navajo. In English, the effect is entirely probabilistic. In Dutch, a conventional or obligatory line is drawn fairly far down the line, to exclude inanimate instruments as subjects. And in Navajo, the entire hierarchy is obligatory (Perkins, 1979, see Reference note 2).

7. Parallel versus serial processing

The competition model makes no assumptions that require serial processing of two or more simultaneously-available pieces of information. Insofar as the model involves mapping coalitions of forms onto coalitions of functions (and vice-versa), we assume parallel activation similar to Selfridge’s Pandemonium model (Selfridge, 1959) in which interpretive ‘demons’ of different weights are activated as soon as their signal comes in. To illustrate, take the simple sentence ‘The boy picks the flower’ in a comprehension task. As soon as the first noun phrase is registered, the listener may simultaneously expect (1) that ‘boy’ will be the subject of a verb that follows, and
that the verb will be compatible with an animate subject. Both expectations are confirmed when the verb 'picks' is processed, and new expectations are set up that (1) the noun phrase that follows will be the object of the transitive verb 'picks', and (2) that noun phrase will be semantically consistent with 'the act of picking'. In this regard, the model is compatible with findings by Marslen-Wilson and Tyler (1980) regarding parallel processing in on-line sentence comprehension, and with proposals by McNeill (1979) for parallel planning in sentence production. On the other hand, the model is not compatible with proposals by Forster (1979), among others, concerning serial application of syntactic and lexical processing.

8. Reaction time

In psycholinguistic research, reaction time data have been used to estimate several different parameters, including number of operations executed (e.g., Miller, 1962), number of elements processed (Sternberg, 1967), and the depth to which processing is carried (Craik and Lockhart, 1972). In many studies, there is a general assumption that more information requires more processing time. This is particularly true of models that assume the existence of separate components, applied in serial order.

An alternative is to view reaction time as a function of the clarity and accessibility of a response, and the speed with which alternative responses can be eliminated. For example, phenomena such as the semantic distance effect (Potts, 1972) or semantic priming (Meyer and Schvaneveldt, 1976) can be viewed as evidence that speed of response is greatest when the accuracy of the response is beyond question. In tests of prototype theory, Rosch (1977) has demonstrated that latency to identify the prototypic member of a category was faster than latency to identify a more peripheral member, i.e., a member with fewer of the features that define the 'best instance'. Hence even though the prototype has 'more' features to be processed, it is accessed more quickly. Such findings are compatible with assumptions of parallel processing, in which underlying representations are activated as soon as the appropriate signals are available.

Insofar as the competition model postulates that grammatical categories are organized around prototypic members (at both the surface and functional level), reaction time predictions from prototype theory would apply. Sentences with 'good' subjects should be analyzed more rapidly and more consistently than sentences with less prototypic subjects and/or sentences with competing interpretations.
9. Non-autonomy of components

In this model, we have differentiated among separate sources of information, i.e., word order versus lexical items at the surface level, topic versus agent at the functional level. However, nothing in the model requires us to group different kinds of information together into separate components (e.g., semantic versus syntactic). In fact, the processes of convergence and competition as we have described them here would be best served by an information-processor without firm boundaries between components, permitting rapid and simultaneous interactions of vectors.

Research strategies and predictions

In a model in which coalitions of functions are mapped onto coalitions of surface forms, data from spontaneous speech may yield very limited information about the underlying system of vector weights that make up those coalitions. Insofar as these alliances evolved to meet high-probability situations, violations of the prototypic situation are (by definition) unlikely to occur. Hence the most efficient research strategy for elaborating the internal structure of the system is to present native speakers with unnatural situations, where forms and functions that usually occur together are placed in competition.

The situation is analogous to the study of perceptual organizing principles by Gestalt psychologists (e.g., Wertheimer, 1923). Because perceptual organizing principles evolved to meet the organization of the real world, it is difficult to specify exactly how they operate in natural situations. As a result, Gestalt psychologists have traditionally relied on the use of visual illusions, which set different organizing tendencies into competition with one another and permit alternative interpretations to emerge.

The use of deviant and/or ambiguous stimuli in sentence processing experiments is not new (e.g., Garrett, 1970). However, studies from the point of view of generative grammar have typically used probabilistic data to demonstrate the existence and interaction of discrete, determinate rules or structures. In research within the competition model, variability that passes as 'noise' in other studies is used to estimate strengths or probabilities that are an inherent part of the mapping system. In a determinate model, two obligatory 'rules' should each produce 100% performance (although there may be noise in the data). Since both rules are determinate, there is no reason why their effects should add. And if the two rules are set in conflict, nothing in the competence model itself predicts which of the two rules would 'win'.
By contrast, the competition model offers the following predictions concerning converging and competing configurations of cues.

A. Convergence

The most rapid reaction times, and most consistent decisions across a group of subjects and within individual subjects, will be produced by a 'prototypic' or 'best input'—a sentence in which all surface sources of information converge on the same interpretation. Sentences with less information (e.g., word order operating alone in reversible sentences, without animacy contrasts) will take longer to process than sentences with more information in the same direction (e.g., the same word order with converging contrasts in animacy). Differences in reaction times and consistency of response for two vectors (e.g., word order versus animacy) can be used to assess their respective strengths.

B. Competition

In competition situations (e.g., word order versus animacy), decisions and reaction times will be a function of the relative weights of competing sources of information. A strong vector competing with a weak one should yield fairly consistent and rapid data, albeit slower than in situations of convergence. Two vectors of roughly equal weights (both weak, or both strong) comprise a 'stand off' which should take longer to resolve, with less consistent interpretations across the group.

C. Conspiracies

When more than two variables are manipulated in the same sentence, it is possible for a 'conspiracy of weak vectors' to emerge. That is, relatively weak sources of information may ally with the 'loser' of a close race to throw the outcome in a different direction, or at least to delay the decision significantly. For example, in a situation where order and animacy are in competition, alliance with contrastive stress may change the outcome.

There are several studies with children using a research strategy compatible with this model. For example, Strohner and Nelson (1974) examined English children's comprehension of non-reversible sentences in which word order competed with animacy (as in 'The baby feeds the mother'). Before the age of 4, English children seem to rely on semantic strategies (see also Bever, 1970). Between 4 and 5, there is a gradual switch to reliance on word order. Slobin and Bever (see Reference note 3) have conducted a cross-
linguistic study of the development of sentence comprehension from 2.3 to 4.4 years of age, in English, Serbo-Croatian, Italian and Turkish. Only reversible sentences were used, so there is no contrast here between semantic and syntactic strategies. However, in the two word order languages (English and Italian), children were presented with converging and competing combinations of word order and contrastive stress. In the two inflectional languages (Serbo-Croatian and Turkish), the sentences included converging and competing combinations of word order and case inflections. In the word order languages, order developed faster and was used far more consistently than contrastive stress. However, Italian children did make greater use of stress information than English children. In the inflectional languages, case inflections were learned earlier and applied more consistently than word order. However, Serbo-Croatian children made greater use of order (and less use of their very difficult case system) than Turkish children. Finally, the most rapid acquisition of all took place in Turkish: Turkish children had acquired almost all of their case system by age 2 1/2, well before case acquisition in Serbo-Croatian or stability of word order cues in either of the word order languages. In the Slobin and Bever data, and in the results presented by Strohner and Nelson, the relative determining force of different kinds of cues shifts gradually over time. Such facts could be captured within a model that specifies weighted probabilities for cues, so that apparent qualitative shifts from one strategy to another actually reflect quantitative shifts in the strength of strategies that exist side-by-side for several years.

In research with adults, we are aware of only one previous sentence processing study which tests the competition/convergence principles directly. This is a study of sentence interpretation in Hebrew by Frankel and Arbel (1980, see Reference note 1). In Hebrew, several different kinds of surface cues mark the subject of a main verb. These include number and gender agreement between subject and verb, an accusative affix on the object of transitive verbs, and preverbal position in NVN sentences (although SVO ordering is not obligatory, and alternative word orders are quite common). Frankel and Arbel presented adult listeners with NVN sentences in which word order converged or competed with two of the three kinds of possible non-order cues: number agreement, and gender agreement. Dependent variables were percent choice of first noun as subject, and reaction time. Although Frankel and Arbel do not use our terminology, their results can be quite directly expressed in the terms that we have just outlined.

Regarding the convergence principle, choice of first noun in Hebrew was 100% when all three cues converged on the same interpretation. However, different combinations of cues favoring the first noun suggest that the effects of the three kinds of information are interactive. Word order only
produced 95.8% first noun choice. When gender only was added to the first noun, the average increased to 99.2%; when number only was added, the average was 98.6%. These differences, though small, were significant. Reaction time data showed the same "degree of convergence" relationships.

The competition items produced more dramatic results. When the order cue conflicted with number agreement on the second noun, then listeners chose the first noun only 7.5% of the time. A conflict between word order and gender produced first noun choice of only 10.5%. In other words, the two agreement cues are much stronger than word order in Hebrew.

Finally, the conspiracy principle was tested here with three "two against one" combinations: word order versus number and gender, word order plus gender versus number, and word order plus number versus gender. As we would expect, the first of these conspiracies (word order versus number and gender) yielded only 2.4% first noun. In other words, when the two agreement markers gang up against word order, then the order cue is overwhelmed. Allied with gender only, word order fared better but still lost the battle: 15.4% first noun choice. When order allied with number, things turned around entirely: 89.6% first noun choice. In other words, number agreement made a much more powerful ally than gender. As we would predict, these two latter conspiracies comprised a stand-off: the slowest reaction times of all were obtained with the order/gender alliance (two losers joining forces). The next slowest were obtained with the order/number alliance (with gender putting up a creditable fight).

Frankel and Arbel note that this technique could be applied not only to differences among grammatical rules, but to combinations of syntactic, pragmatic and semantic cues. That is precisely what we have done in the present study. Italian and English adults are given a series of simple sentences with two nouns plus a verb. Unlike the Frankel and Arbel study, all possible word order combinations are used. These are set in competition and convergence with semantic information (i.e., reversible sentences, and non-reversible sentences that agree or disagree with word order information). In addition, two kinds of surface cues that are often associated with pragmatic information are also manipulated: contrastive stress, and topicalization or givenness/newness. These are also placed in competition or convergence with both word order and lexical input.

Because of the complexity of the competition system the design of the experiment is also complex. In addition, ancillary analyses plus an additional experiment were required to control for some alternative interpretations. Thus the results section is necessarily lengthy and difficult. We have tried to hold details to a minimum, and included as many summary statements as possible throughout. But we must request patience from the reader.
Method

Subjects

60 middle-class adult subjects participated in the experiment, 30 native speakers of Italian and 30 native speakers of English. The majority of the subjects were university students in their 20’s, although some post-graduate adults were included in the Italian sample. All English participants were residents of the Denver/Boulder area. All Italians were residents of Rome, speaking a dialect typical of educated residents of that city.

Materials

A set of 18 toy plastic animals, 1 1/2—2 inches high, and 9 small inanimate objects of roughly equivalent size, were used to accompany the presentation of test sentences (see Table 1). Although these visual stimuli were not necessary for the adult study, we planned to carry out an analogous study with children in both language groups. Hence the same animals and objects were included here to insure that whatever biases they introduced were the same for adults and children. Responses were audiorecorded with portable cassette recorders for later transcription. A stopwatch was used to take reaction time data from the audiotape.

Test sentences

There was a total of 81 test sentences for each subject, each comprised of a verb in the third person singular, and two common nouns with definite or indefinite articles. There were 27 nouns (the 18 animals and 9 inanimate objects listed in Table 1) and 9 verbs (also listed in Table 1). Because individual lexical items might be biased in terms of the plausibility of their combinations (e.g., lions are perceived as more ferocious and active than sheep), combinations of nouns and verbs were assigned randomly to each sentence type, for each subject. Hence specific lexical effects were completely homogenized across subjects and experimental conditions.

The factors manipulated as independent variables include language group, three levels of word order (NVN, NNV, VNN), three levels of contrastive stress (no stress, stress on the first noun, stress on the second noun), three levels of animacy (both animate, first noun animate and second inanimate, first noun inanimate and second animate), and three levels of topicalization (no topicalized element, first noun topicalized, second noun topicalized). The topic manipulation consisted of introducing one of the two objects
Table 1. *Object names and verbs used in Experiment I*

<table>
<thead>
<tr>
<th>Animate objects</th>
<th>Inanimate objects</th>
<th>Verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italian</td>
<td>English</td>
<td>Italian</td>
</tr>
<tr>
<td>gatto</td>
<td>cat</td>
<td>mangia</td>
</tr>
<tr>
<td>cavallo</td>
<td>horse</td>
<td>annusa</td>
</tr>
<tr>
<td>cammello</td>
<td>camel</td>
<td>lecca</td>
</tr>
<tr>
<td>giraffa</td>
<td>giraffe</td>
<td>bacia</td>
</tr>
<tr>
<td>capra</td>
<td>goat</td>
<td>morde</td>
</tr>
<tr>
<td>vitellino</td>
<td>calf</td>
<td>guarda</td>
</tr>
<tr>
<td>scimmia</td>
<td>monkey</td>
<td>carezza</td>
</tr>
<tr>
<td>orso</td>
<td>bear</td>
<td>afferra</td>
</tr>
<tr>
<td>malalino</td>
<td>piglet</td>
<td>saluta</td>
</tr>
<tr>
<td>mucca</td>
<td>cow</td>
<td></td>
</tr>
<tr>
<td>cerbiatto</td>
<td>deer</td>
<td></td>
</tr>
<tr>
<td>zebra</td>
<td>zebra</td>
<td></td>
</tr>
<tr>
<td>pecora</td>
<td>sheep</td>
<td></td>
</tr>
<tr>
<td>agnellino</td>
<td>lamb</td>
<td></td>
</tr>
<tr>
<td>tartaruga</td>
<td>turtle</td>
<td></td>
</tr>
<tr>
<td>cane</td>
<td>dog</td>
<td></td>
</tr>
<tr>
<td>asinello</td>
<td>donkey</td>
<td></td>
</tr>
<tr>
<td>pesce</td>
<td>fish</td>
<td></td>
</tr>
</tbody>
</table>

first, as follows: 'This is a (cow); now we’re going to talk about this (cow): the (cow) a (horse) (kicks)'. Hence both the preamble and the contrast between definite and indefinite articles marked one of the nouns as given and the other as new. On the items without a topic manipulation, both nouns were presented with definite articles. This is admittedly an artificial topic manipulation, very indirectly related to the natural build-up of theme, givenness and newness in connected discourse. This fact must be kept in mind in interpreting the results.

The five independent variables were orthogonalized, yielding 81 sentences within a 2 (language) × 3 (word order) × 3 (animacy) × 3 (stress) × 3 (topic) design. With these 81 sentences, there was only one unique sentence for each complete order/animacy/stress/topic combination. Hence the power of possible five-way interaction is considerably weakened. However, we felt that this was more feasible than doubling or tripling the number of items across all cells.

The 81 sentences were divided into nine blocks of nine items each, for purposes of the ‘induced introspection’ procedure which we will explain shortly. Each block contained all nine possible order/stress combinations, in one of the three animacy conditions and one of the three topic condi-
tions. The order of sentences within blocks, and the order of the nine blocks themselves, was randomized individually for each subject.

Procedure

There were two parts of the experiment: sentence interpretation, and an 'induced introspection' procedure in which we obtained subjects' intuitions about the factors that influenced their decisions. The reason for the second procedure was primarily methodological. As Gleitman and Gleitman (1970) have noted, individuals may vary considerably in the extent to which they interpret an experiment as a 'test of common sense' versus a test involving grammar. It is possible that one could obtain differences between Italians and Americans in their use of semantic versus syntactic information, not because of true language differences but because of cultural differences in interpreting the goals of the experiment. To minimize this possibility, subjects were informed explicitly, in a standardized and unbiased way, of the factors which were 'fair game' for consideration in sentence interpretation. This involved, first, listing in advance the factors that would be varied: word order, presence or absence of emphatic stress, whether or not both of the nouns were animate, whether or not one of the nouns was introduced as the topic of the discussion, and finally, different kinds of verbs. Although the plausibility of different animate/verb/inanimate combinations was not varied systematically it seemed quite possible that subjects would also use that information in making their decisions. Ideally, it would have been desirable to ask for the subject's intuitions after every test sentence. However, with 81 sentences that seemed most impractical. Instead, the nine blocks of items were designed to reflect the possible combinations that listeners could use in making up their minds. By separating out blocks with and without topic and animacy contrasts, it was possible to assess the listeners' intuitions about the importance of grammatical information with and without competition from these semantic-pragmatic variables. Later we will consider some information from a subsidiary study without this procedure, ruling out the possibility that results are artificially skewed by this unusual set of instructions.

Subjects were tested individually by female graduate students; native speakers of Italian tested the Italian subjects, and native speakers of English tested the American subjects. The instructions and test sentences were all read aloud to each subject in a standardized form (equivalent translations for the two respective language groups). The experimenter explained the cross-linguistic nature of the experiment, and the fact that a subsequent study would examine the same comprehension test with children learning
different languages. This explanation was included to justify (a) the presentation of the small animals and objects with each test sentence, and (b) the fact that so many of the items might seem 'odd' in the listener's own language. Great care was taken not to bias subjects toward a syntactic versus a semantic strategy by essentially asking for both:

'I will read you a series of very simple sentences. After I read each sentence, you will have to interpret it: you should tell me which one of the two nouns in the sentence is (the subject of the sentence), that is, (the one who does the action).'

Half the subjects were given these instructions with 'subject' first and 'one who does the action' second; half received the same instructions in the opposite order. The experimenters also stressed that different ways to interpret the sentences were possible, and that there was no 'right' answer that we had decided on in advance. At that point, the list of factors that would be varied from one sentence to another were listed and explained (i.e., 'the order of the words', topic, animacy, contrastive or emphatic stress, and different kinds of verbs). These factors were always listed in one of several randomized orders, which were in turn assigned randomly to subjects, to avoid indicating that any one of them was 'the' important variable. The experimenter also explained in advance that she would stop periodically and ask the subject to explain how he had reached his decisions so far.

After each block of nine sentences, the experimenter stopped and listed the factors that had been varied (again, in one of several randomized orders to avoid biasing any one of them). Subjects were then asked to rank order these factors from the one(s) that played the most important role in interpretation, on down to factors that were not taken into consideration at all. Tie ranks were permitted, so that in principle none of the rankings were mutually exclusive. It appears that none of these educated adult subjects had any difficulty with this part of the task.

The fact that test sentences were presented orally means that there was no absolute mechanical control over pauses, sound levels in contrastive stress, etc. However, because each subject received a completely different random configuration of nouns and verbs, in random orders, such mechanical control was impractical. Also, since subjects had already been told explicitly that stress would be varied, they obviously expected to hear it. The sentences were read in as standard a fashion as possible, with very clear distinctions between the default stress versus contrastive stress sentences. Slight differences within these three stress levels (no stress, stress on first noun, stress on second) are unlikely to have affected results in any systematic way.
Reaction times were recorded from taped responses with a hand stopwatch which was started manually at the voice-offset of the test sentence and stopped at the listeners' pronunciation of the chosen noun. While this is a crude measure by the standards of some types of reaction time studies, it was appropriate given the rather large individual variation between items and between subjects that resulted in this task. The relevant range is from 320 to 28,270 milliseconds.

Results and discussion

We will begin by presenting the results of the two major analyses of variance, on the respective choice and latency measures. Since the two dependent variables are complementary in many interesting respects, it may be most useful to present the results side-by-side, comparing choice and latency effects at each respective level of the analysis. Some ancillary correlational analyses are also presented, comparing choice and latency directly across sentence types. Then we will turn to the results of the induced introspection procedure, including correlations between these subjective judgments and actual behavior in the experiment. Finally, to rule out some alternative interpretations of the main experiment, we will consider some subsidiary analyses, plus a control experiment in English without the induced introspection procedure.

Analyses of variance on choice and latency

Each of the two dependent variables was analyzed within a 2 X 3 X 3 X 3 X 3 analysis of variance (language X word order X animacy X stress X topic), with language as a between-subjects variable and the other four factors as repeated measures. The raw latency data showed considerable variation between individuals in mean reaction time, variation which could have masked experimental effects. For this reason, the analysis of variance was carried out on z-score transformations of each subject’s raw latency scores, so that each subject is actually compared against his own mean reaction time across experimental conditions.

Table 2 summarizes significant results from the choice ANOVA; Table 3 summarizes significant results from the latency ANOVA. In those tables, we have included some information about the amount of variance accounted for by each main effect and interaction. Although it is not traditional to include such statistics, we feel that information concerning the power of these effects is just as important, and in some instances more illuminating.
Table 2. *Summary of significant effects (p < 0.05) from the analysis of variance on percent choice of first noun*

<table>
<thead>
<tr>
<th>Main effects &amp; interactions</th>
<th>d.f.</th>
<th>F-ratio</th>
<th>p &lt;</th>
<th>Percent of variance accounted for</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>1,58</td>
<td>70.49</td>
<td>0.0000001</td>
<td>4.00%</td>
</tr>
<tr>
<td>Word Order</td>
<td>2,116</td>
<td>255.40</td>
<td>0.0000001</td>
<td>20.00%</td>
</tr>
<tr>
<td>Animacy</td>
<td>2,116</td>
<td>173.04</td>
<td>0.0000001</td>
<td>16.00%</td>
</tr>
<tr>
<td>Stress</td>
<td>2,116</td>
<td>3.37</td>
<td>0.04</td>
<td>0.10%</td>
</tr>
<tr>
<td>Topic</td>
<td>2,116</td>
<td>8.46</td>
<td>0.0006</td>
<td>0.30%</td>
</tr>
<tr>
<td><strong>2-Way Interactions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language x Word Order</td>
<td>2,116</td>
<td>102.96</td>
<td>0.0000001</td>
<td>0.80%</td>
</tr>
<tr>
<td>Language x Animacy</td>
<td>2,116</td>
<td>58.73</td>
<td>0.0000001</td>
<td>0.50%</td>
</tr>
<tr>
<td>Language x Stress</td>
<td>2,116</td>
<td>3.36</td>
<td>0.04</td>
<td>0.10%</td>
</tr>
<tr>
<td>Word Order x Animacy</td>
<td>4,232</td>
<td>16.77</td>
<td>0.0000001</td>
<td>0.80%</td>
</tr>
<tr>
<td>Word Order x Stress</td>
<td>4,232</td>
<td>3.34</td>
<td>0.02</td>
<td>0.07%</td>
</tr>
<tr>
<td>Word Order x Topic</td>
<td>4,232</td>
<td>4.16</td>
<td>0.0004</td>
<td>0.10%</td>
</tr>
<tr>
<td><strong>3-Way Interactions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language x Word Order x Animacy</td>
<td>4,232</td>
<td>2.47</td>
<td>0.05</td>
<td>0.10%</td>
</tr>
<tr>
<td>Language x Word Order x Topic</td>
<td>4,232</td>
<td>2.51</td>
<td>0.05</td>
<td>0.06%</td>
</tr>
<tr>
<td>Language x Animacy x Stress</td>
<td>4,232</td>
<td>3.21</td>
<td>0.02</td>
<td>0.80%</td>
</tr>
<tr>
<td><strong>4-Way Interactions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language x Word Order x Animacy x Stress</td>
<td>8,464</td>
<td>3.63</td>
<td>0.0007</td>
<td>0.10%</td>
</tr>
</tbody>
</table>

1 Total sum of squares = 1214.99. Eta-square statistic for variance accounted for by experimental variables = 69%.

...than p-values documenting the reliability of the same main effects and interactions. In this regard, an eta-square statistic was calculated for each ANOVA, showing the amount of variance accounted for by all the experimental effects and interactions taken together. In the choice ANOVA, experimental effects accounted for 69% of the total variance; in the latency ANOVA, experimental effects accounted for only 11% of the variance. In other words, there was considerably more noise in the latency data; hence the power of each individual effect is correspondingly reduced.

Because there were so many significant interactions in both analyses, a detailed examination of each one would take considerable space. Also, since there was a significant interaction at the four-way level, a detailed account of each two- and three-way effect could also be misleading. For these reasons, we will shorten the presentation as follows. First, we will briefly characterize the nature and direction of the main effects in both analyses. Then we will discuss the *language x word order, language x animacy*, and
Table 3. Summary of significant effects (p < 0.06) from the analysis of variance on latency scores

<table>
<thead>
<tr>
<th>Main effects &amp; interactions</th>
<th>d.f.</th>
<th>F-ratio</th>
<th>p</th>
<th>Percent of variance accounted for</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Order</td>
<td>2,116</td>
<td>38.28</td>
<td>0.0000001</td>
<td>2.00%</td>
</tr>
<tr>
<td>Animacy</td>
<td>2,116</td>
<td>9.48</td>
<td>0.0004</td>
<td>1.00%</td>
</tr>
<tr>
<td>2-way interactions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language x Word Order</td>
<td>2,116</td>
<td>6.69</td>
<td>0.003</td>
<td>0.40%</td>
</tr>
<tr>
<td>Language x Animacy</td>
<td>2,116</td>
<td>17.36</td>
<td>0.000001</td>
<td>2.00%</td>
</tr>
<tr>
<td>Word Order x Animacy</td>
<td>4,232</td>
<td>18.57</td>
<td>0.000001</td>
<td>2.00%</td>
</tr>
<tr>
<td>Word Order x Stress</td>
<td>4,232</td>
<td>2.28</td>
<td>0.06</td>
<td>0.10%</td>
</tr>
<tr>
<td>3-way interactions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language x Word Order x Animacy</td>
<td>4,232</td>
<td>5.58</td>
<td>0.0005</td>
<td>0.40%</td>
</tr>
<tr>
<td>Language x Word Order x Stress</td>
<td>4,232</td>
<td>2.83</td>
<td>0.03</td>
<td>0.20%</td>
</tr>
<tr>
<td>4-way interactions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language x Word Order x Animacy x Stress</td>
<td>8,454</td>
<td>2.37</td>
<td>0.02</td>
<td>0.30%</td>
</tr>
</tbody>
</table>

1 Total sum of squares = 4800.01. Eta square statistic for variance accounted for by experimental variables = 11%.

language x word order x animacy relationships in a separate section. Then all the relationships with topic will be considered together. Finally, we will look at the effect of stress within the significant four-way interaction of language x word order x animacy x stress.

Main effects in the choice ANOVA

All five of the factors in the choice analysis reached significance. Of these five, word order and animacy were not only quite reliable (p < 0.0000001) but also very large. When considered separately, they accounted for 20% and 16% of the variance, respectively. By contrast, topic and stress effects were reliable but small: stress alone accounted for 0.1% of the variance, topic for 0.3%. In general, then, it seems fair to conclude that word order and animacy are the ‘primary vectors’ determining sentence interpretation in this experiment, while topic and stress are ‘secondary vectors’ that operate primarily in interaction with the larger effects. Since all these main effects also enter into complex interactions, they should be interpreted with caution. However, the following summary statements seem fair.
Main effect of language
Italians chose the first noun more often than Americans (60% versus 41%). However, as we shall see shortly, this is not because Italian use of order is more systematic. In fact, the opposite is true.

Main effect of word order
As we would expect summing across two SVO languages, the first noun was chosen more often in NVN items (82%), compared with 31% for VNN and 37% for NNV. Thus subjects tend to choose the noun preceding the verb as the actor and the noun following the verb as the object.

Main effect of animacy
Summing across language groups, the first noun is selected more often in sentences with an animate noun first and an inanimate second (74%), compared with 25% first noun choice in inanimate/animate orders and 48% for the reversible animate/animate items.

Main effect of topic
There was a slight but significant tendency for the topicalization manipulation to influence choice. The first noun was chosen 54% of the time when it was topicalized, compared with 51% when neither element was topicalized, and 46% when the second was topicalized.

Main effect of stress
A very simple prediction regarding the effect of stress would be that listeners tend to interpret a stressed element as new, and hence more likely to be the object. But matters were not so simple. The first noun was chosen 52% of the time when there was no contrastive stress at all, compared with 48% when the first element was stressed and 50% when the second element was stressed. Further interpretation of the stress data must be delayed until we examine the complete four-way interactions of stress with configurations or order/animacy between languages.

Main effects in the latency ANOVA
There were only two significant main effects in the latency ANOVA: word order and animacy. The word order effect accounts for only 2% of the vari-
ance, but since the latency analysis as a whole produced an eta-square of only 11%, this is still a relatively large effect and certainly a reliable one (p < 0.0000001). Animacy accounted for 1% of the variance, significant at p < 0.0004. Again, interpretations of these main effects are offered with caution, since they participated in complex interactions. But the following summary statements are offered.

**Main effect of word order**
NVN sentences were resolved significantly faster than non-canonical VNN and NNV orders—again, in the predicted direction for two SVO languages. The mean z-scores (where negative scores indicate more rapid reaction times) were -0.22 for NVN, +0.12 for VNN and +0.10 for NNV.

**Main effect of animacy**
Mean reaction times were slowest in reversible sentences without an animacy contrast (+0.15). For sentences in an animate/inanimate order, mean latencies were -0.09; for sentences in an inanimate/animate order, the mean was -0.07. In other words, summing across languages and order/stress/topic configurations, decisions are reached more quickly when animacy information contributes to sentence interpretation. However, as we shall see, this effect is due almost entirely to the Italians.

The absence of topic and stress main effects on latency corresponds to our interpretation of the choice data: these are secondary vectors that affect sentence processing primarily through interaction with the primary vectors of order and animacy. That is, topic and stress may either slow or speed reaction times, depending on the alliances that they form in particular order/animacy configurations.

**Language \times word order \times animacy in both analyses**
We have already noted that word order and animacy are the two major determinants of sentence interpretation in this experiment. However, the relative weights of these two vectors are completely opposite in Italian versus English. Summing across the two languages, order and animacy accounted, respectively, for 20% and 16% of the variance in choice and 2% and 1% of the variance in latency. However, if we treat the Italian and English samples as separate experiments, the corresponding figures look very different. In the English sample, word order accounts for 51% of the choice variance, com-
pared with only 3% for animacy; in the Italian sample, word order accounts for only 4% of the choice variance, compared with 42% for animacy. In the latency ANOVA, English figures are 62% for word order and 6% for animacy; Italian figures are 10% for word order and 59% for animacy. In other words, Americans rely on order information while Italians rely on the lexical information providing the animacy contrast.

Before looking in detail at the language × word order × animacy interaction, let us first see how these mirror-image effects operate at the two-way level.

Language × word order

Figure 1 illustrates this interaction in the choice data; Figure 2 illustrates the same interaction in latency. Starting with the choice data, the first difference that emerges is the contrast between American and Italian per-
performance on standard NVN items. Americans choose the first noun in NVN an average of 92\% of the time; for Italians, the same figure is only 71\%. If we view 'SVO-ness' as a dimension rather than a discrete classification, we might argue that English is a stronger SVO language than Italian.

There is also a difference in the strategies Americans and Italians use to interpret the two non-canonical orders. We suggested earlier that pragmatic reorderings are structurally constrained in informal English: left-dislocations (OSV) and right-dislocations (VOS) both appear in informal conversation, but SOV and VSO do not. By contrast, all word orders can and do occur in informal Italian. These facts about English and Italian production are
mirrored in the comprehension results obtained here. Specifically, English listeners make very consistent use of what we will call a second noun strategy, selecting the second noun in VNN and NNV an average of 85% of the time (hence yielding VOS and OSV interpretations). There is no such consistent strategy in Italian: 48% first noun on VNN, and 60% first noun on NNV. Indeed, the slight SOV bias on the NNV items is in the opposite direction from their American counterparts.

Another way to look at these data is to ignore direction, examining only the consistency of these strategies. A direct measure of consistency can be obtained by subtracting mean choice scores from 50%, taking the absolute value. By this procedure, we obtain consistency measures in English of 0.42 for NVN, and 0.35 for the two alternative order. In other words, there is a surprisingly small difference in English between the strength of the ‘grammatical’ SVO interpretation and the strength of the ‘semi-grammatical’ VOS and OSV interpretations. The corresponding consistency measures for Italian are 0.21 for NVN, 0.02 for VNN and 0.10 for NNV.

For the sake of brevity, we will refer to the English tendencies as first- and second-noun strategies throughout. We should note, however, that these apparent holistic strategies might actually be epiphenomenal of interactions among several ‘local’ interpretative principles. For example, the NVN interpretation may involve two strategies: preverbal noun = subject; post-verbal noun = object. On NNV, the preverbal noun strategy might be all that applies (since nothing follows the verb). On VNN, the post-verbal strategy might produce VGS interpretations (since nothing precedes the verb). Very different principles might guide Italian interpretations. For example, the slight bias in favor of SOV could be affected by the fact that clitic objects obligatorily occur directly in preverbal position. Since we cannot decide among these competing alternatives here, we will use the more holistic terms as short-hand.

Turning to the latency data, we obtain a rather different view of these results. As illustrated in Figure 2, both Americans and Italians took considerably longer to make decisions on non-standard word orders. This is not surprising in the Italian case, since they were also clearly ‘stuck for an answer’. But for the Americans, the latency data tells us something new: consistency in output can belie differences in processing. That is, the second noun strategy is somehow a ‘weaker’ even though it is quite consistent; VOS and OSV are more difficult to access or derive, compared with SVO.

Figure 2 also suggests that Italians are faster than Americans on two of the three word orders. However, this tendency is an artifact created by summing across animacy conditions, as we shall see in examining the language × animacy interaction.
Language × animacy

Figure 3 illustrates this two-way interaction in the choice ANOVA; Figure 4 illustrates the same effect in latency.

Starting with the choice data, the reversible sentences reflect nothing other than the sum of the word order tendencies just described: Italians choose the first noun 66% of the time (a relatively strong SVO bias + a relatively weak SOV bias + random levels on VNN); Americans choose the first noun 38% of the time (SVO + OSV + VOS). However, when the animacy contrast is added, the divergence between English and Italian is quite dramatic. On items with an AI ordering (animate/inanimate), Italians choose the first noun 95% of the time compared with 53% for Americans. On items with an IA ordering, Italians choose the first noun only 8% of the time, compared with 31% for Americans. English listeners are affected by the animacy manipulation, but to a far lesser degree than their Italian counterparts.

Figure 3. Percent choice of the first noun as actor: Language by animacy order.
Turning to the latency data, we find still further evidence that animacy is the favored source of information for Italians. In reversible items, where sentence interpretation is determined almost entirely by word order, Italian reaction times are extremely slow: +0.32 compared with −0.03 for Americans, averaged across order/stress/topic configurations. However, on non-reversible items the Italians are much faster than Americans: −0.16 versus +0.01 (averaging AI and IA items). This gives us an important new piece of information: it is not simply the case that Italians prefer to use animacy, but that they have also more difficulty using word order when that is the only source of information available.
Language × word order × animacy

The cell means for this three-way interaction, in both the choice and latency analyses, are presented in Table 4. We also included the choice consistency measure (percent choice of first noun — 50% for all nine possible order/animacy combinations, in each language.

Table 4. Order and animacy configurations: choice and latency data

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Choice</td>
<td>Choice consistency</td>
</tr>
<tr>
<td>NVN:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td>0.96</td>
<td>0.46</td>
</tr>
<tr>
<td>A/I</td>
<td>0.98</td>
<td>0.48</td>
</tr>
<tr>
<td>I/A</td>
<td>0.82</td>
<td>0.32</td>
</tr>
<tr>
<td>VNN:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td>0.09</td>
<td>0.41</td>
</tr>
<tr>
<td>A/I</td>
<td>0.32</td>
<td>0.18</td>
</tr>
<tr>
<td>I/A</td>
<td>0.05</td>
<td>0.45</td>
</tr>
<tr>
<td>NNV:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td>0.10</td>
<td>0.40</td>
</tr>
<tr>
<td>A/I</td>
<td>0.28</td>
<td>0.22</td>
</tr>
<tr>
<td>I/A</td>
<td>0.07</td>
<td>0.43</td>
</tr>
</tbody>
</table>

R_{xy} of consistency to latency = -0.82  
R_{xy} of consistency to latency = -0.92

In this interaction, we can see the reciprocal influence of order and animacy on one another. A very general summary of this three-way effect is that animacy has a greater impact on the interpretation of non-standard word orders. However, we can derive much more detailed information by applying the convergence/competition analysis proposed earlier.

Convergence

The ‘best input’—at this level, an item in which word order and animacy point to the same interpretation—should produce faster and more consistent decisions than the same word order acting alone, or the same word
order competing with animacy. In English, the 'best' or preferred readings for each of the three word order types should be AVI (SVO + animacy versus OVS), IAV (OSV + animacy versus SOV), and VIA (VOS + animacy versus VSO). For Italian, the 'best' NNV should also be AVI. The best NNV is in the opposite direction from the American preference, and hence should be AIV. The best VNN is difficult to determine, since performance was so close to chance, but since decisions were very slightly in favor of VOS, we will assume that the best case is VIA. This situation yields the following 12 predictions (where > indicates 'better than'):

<table>
<thead>
<tr>
<th>Convergence &gt; Order Alone</th>
<th>English</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVI &gt; AVA</td>
<td>AVI &gt; AVA</td>
<td></td>
</tr>
<tr>
<td>IAV &gt; AAV</td>
<td>AIV &gt; AAV</td>
<td></td>
</tr>
<tr>
<td>VIA &gt; VAA</td>
<td>VIA &gt; VAA</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Convergence &gt; Competition</th>
<th>English</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVI &gt; IVA</td>
<td>AVI &gt; IVA</td>
<td></td>
</tr>
<tr>
<td>IAV &gt; IAV</td>
<td>AIV &gt; IAV</td>
<td></td>
</tr>
<tr>
<td>VIA &gt; VAI</td>
<td>VIA &gt; VAI</td>
<td></td>
</tr>
</tbody>
</table>

Applying these same 12 predictions to both the reaction time and choice data, we have a total of 24 tests of the convergence principle. As can be seen from Table 4, all 24 predictions are supported.

We can also use these data to estimate the relative strength of word orders in Italian versus English, by looking only at the reversible items. In the consistency data, the spread among the reversible items for the three order types was fairly small in English: 0.46 for AVA, 0.40 for AAV and 0.41 for VAA. Hence if we estimate relative strengths from the outcome alone, non-canonicals are not very different from the standard SVO. However, the corresponding latency figures are -0.33 for AVA, +0.19 for AAV, and +0.07 for VAA. We may conclude that VOS is slightly stronger than SOV, but that both are considerably weaker than canonical word order in English.

The Italian consistency data yields a much more distinct differentiation among the order types: 0.34 for AVA, 0.18 for AAV, and a negligible 0.05 for VAA. The same hierarchy is reflected in the latency data: +0.05 for AVA, +0.32 for AAV, and +0.61 for VAA.

There is nothing surprising about the finding that SVO sentences are 'better' in SVO languages. What is somewhat more surprising is the comparison between the two languages. On the AVA items, when both groups were relying exclusively on an SVO strategy, results were very different: Italians were less consistent (84% first noun compared with 95% in English) and much slower (+0.05 compared with 0.33). This confirms the picture that has been building up so far: Italian is less of an SVO language than English, with
or without animacy contrasts. In fact, Americans were more consistent and only slightly slower in their use of VOS and OSV strategies than Italians were with SVO.

**Competition**

As noted earlier, in both the standard and non-standard word order, animacy always 'wins' in Italian and order always 'wins' in English—although (as we would expect) the margin of those respective victories is smaller in the maximum competition situation, IVA. However, when we examine more details within the non-canonical orders, yet another language difference emerges. Americans give much slower and less consistent responses in AIV and VAI than in the AAV and VAA items where word order operates alone. In other words 'competition' is worse than 'weak convergence'. But Italians show the opposite result: their worst performance by far takes place not on the competition items but on the reversibles, AAV and VAA. On IAV and VAI, Italians decide rapidly in favor of animacy. In other words, non-canonical word order biases in Italian are so weak that they really do not provide much competition to animacy.

The last part of our competition/convergence predictions regards conspiracies or 'ganging up' of two or more vectors against another. At the level of language × animacy interactions, this analysis is not applicable. We will examine conspiracy effects later in considering interactions with stress and topic.

**Effects of topicalization**

Topic had no noticeable effect at any level of the latency analysis. In the choice analysis, in addition to the main effect of topic described earlier, there were two significant interactions: word order × topic, and language × word order × topic. These small effects can be summarized very quickly.

**Word order × topic**

The tendency to choose the topicalized element as the agent of the action occurred almost entirely within the non-canonical sentences. This is compatible with the proposal that VNN and NNV orders are pragmatic variants in the service of topicalization. However, it is also compatible with the idea that weak vectors cannot compete against strong ones (e.g. SVO), so that their effects can only show up when word order biases are not very strong.
Language x word order x topic

What this effect basically shows is that Italians are the only ones making significant use of topic information, and this use is further restricted to the non-canonical orders.

As we noted from the outset, this was a particularly artificial topic manipulation, quite dissimilar to the long build-ups of topic that tend to occur in natural discourse (e.g. Karmiloff-Smith, 1979; Keenan and Schieffelin, 1976; Li and Thompson, 1979). It remains to be seen whether a more realistic discourse manipulation might not have a more appreciable influence on sentence interpretation. In this experiment, however, the more credible evidence for pragmatic effects comes from the results for contrastive stress.

Effects of stress on latency and choice

In contrast with the topic manipulation, contrastive stress entered into a large number of significant interactions in both of the analyses of variance (summarized in Tables 2 and 3). In both ANOVAs, the effect of stress reached down to the level of a significant four-way interaction: language x word order x animacy x stress. Cell means for this interaction in both the choice and latency analyses are presented in Table 5, together with means for response consistency (percent first noun 50%).

A very general summary of this four-way interaction is (a) that Italians make more use of stress information than Americans, and (b) that stress has a greater impact on the interpretation of non-canonical orders. To understand exactly how stress is used within particular order/animacy configurations, we can apply some of the convergence and competition principles introduced above.

Since the four-way interaction accounted for only 0.10% of the variance in choice, and 0.30% of the variance in latency, there was no point in applying post-hoc significance tests to all the conceivable pair-wise combinations of 27 item types. Instead, we will take a more cautious and theory-guided approach to interpreting the four-way interaction, restricting ourselves to an examination of how stress affects the extremes, i.e., the 'best' and 'worst' cases for each word order type.

A very global prediction was that contrastive stress would discourage selection of the stressed element as subject, under the assumption that contrastive stress is associated with new information and the subject role is typically associated with given information. This stress/object bias would then be expected to converge or compete with order and animacy informa-
tion in determining both choice and reaction times. However, the situation is not quite so simple. Contrastive stress also serves a more general function, as a signal to the listener that a default or unmarked situation should be suspended. Hence subject or object stress in a standard NVN word order may interfere with processing, suggesting a suspension of the default situation. The ‘best’ stress conditions for an SVO interpretation should be default stress, i.e., no contrastive stress at all. The ‘best’ conditions for interpretation of non-canonical orders should be stress on the inanimate element, in object position. We will go into the analysis with those predictions.

**Convergence: effects of stress on best cases**

The best cases for each order type at the three-way level were AVI, VIA and IAV in English and AVI, VIA and AIV in Italian. How can contrastive stress interact to make these cases ‘even better’? Applying the assumptions that default stress is best in NVN and object stress is best in alternative orders, we arrive at the following predictions (where > indicates ‘better than’):

<table>
<thead>
<tr>
<th>English</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVI &gt; AVI</td>
<td>AVI &gt; AVI</td>
</tr>
<tr>
<td>AVI &gt; AVI</td>
<td>AVI &gt; AVI</td>
</tr>
<tr>
<td>VIA &gt; VIA</td>
<td>VIA &gt; VIA</td>
</tr>
<tr>
<td>VIA &gt; VIA</td>
<td>VIA &gt; VIA</td>
</tr>
<tr>
<td>IAV &gt; IAV</td>
<td>AJV &gt; AIV</td>
</tr>
<tr>
<td>IAV &gt; IAV</td>
<td>AJV &gt; AIV</td>
</tr>
</tbody>
</table>

In the latency data, all 12 of these predictions are confirmed (see Table 5). In the choice data, there are only 5 confirmations in Italian and 2 in English. However, the disconfirming cases involve discrepancies of very few percentage points, so that the larger spread in the latency data probably provides a better test. Given the instability of the data at the four-way level, we should not press our case too far for specific items. However, in very general terms the convergence prediction seems to be met: reaction times to ‘good’ order/animacy configurations are even faster when converging evidence from stress is available.

Of all the above combinations, the ‘best sentence of all’ in both languages is AVI with default stress. The ‘perfect’ SVO produces a 99% first noun choice in both languages, with reaction times of −0.65 in English and −0.56 in Italian. By contrast, the best versions of the other two word order types are significantly slower and less consistent. Remember that Italians are relying primarily on the animacy contrast in all these items, while Amer-

Table 5.  Order/animacy/stress configurations: choice and latency data

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th></th>
<th></th>
<th>Italian</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Choice</td>
<td>Consistency</td>
<td>Latency</td>
<td>Choice</td>
<td>Consistency</td>
<td>Latency</td>
</tr>
<tr>
<td>NVN: A/A</td>
<td>0.96</td>
<td>0.46</td>
<td>-0.40</td>
<td>0.90</td>
<td>0.40</td>
<td>0.04</td>
</tr>
<tr>
<td>A/A</td>
<td>0.96</td>
<td>0.46</td>
<td>-0.38</td>
<td>0.81</td>
<td>0.31</td>
<td>0.00</td>
</tr>
<tr>
<td>A/A</td>
<td>0.98</td>
<td>0.48</td>
<td>-0.22</td>
<td>0.80</td>
<td>0.30</td>
<td>0.10</td>
</tr>
<tr>
<td>A/I</td>
<td>0.99</td>
<td>0.49</td>
<td>-0.65</td>
<td>0.99</td>
<td>0.49</td>
<td>-0.56</td>
</tr>
<tr>
<td>A/I</td>
<td>0.99</td>
<td>0.49</td>
<td>-0.27</td>
<td>0.98</td>
<td>0.48</td>
<td>-0.55</td>
</tr>
<tr>
<td>A/I</td>
<td>0.97</td>
<td>0.47</td>
<td>-0.46</td>
<td>0.97</td>
<td>0.47</td>
<td>-0.25</td>
</tr>
<tr>
<td>I/A</td>
<td>0.82</td>
<td>0.32</td>
<td>-0.13</td>
<td>0.39</td>
<td>0.11</td>
<td>-0.06</td>
</tr>
<tr>
<td>I/A</td>
<td>0.82</td>
<td>0.32</td>
<td>0.00</td>
<td>0.21</td>
<td>0.29</td>
<td>0.00</td>
</tr>
<tr>
<td>I/A</td>
<td>0.81</td>
<td>0.31</td>
<td>-0.14</td>
<td>0.36</td>
<td>0.14</td>
<td>0.04</td>
</tr>
<tr>
<td>VNN: A/A</td>
<td>0.07</td>
<td>0.43</td>
<td>-0.05</td>
<td>0.52</td>
<td>0.02</td>
<td>0.84</td>
</tr>
<tr>
<td>A/A</td>
<td>0.10</td>
<td>0.40</td>
<td>0.20</td>
<td>0.37</td>
<td>0.13</td>
<td>0.50</td>
</tr>
<tr>
<td>A/A</td>
<td>0.10</td>
<td>0.40</td>
<td>0.06</td>
<td>0.47</td>
<td>0.03</td>
<td>0.48</td>
</tr>
<tr>
<td>A/I</td>
<td>0.36</td>
<td>0.14</td>
<td>0.51</td>
<td>0.93</td>
<td>0.43</td>
<td>-0.20</td>
</tr>
<tr>
<td>A/I</td>
<td>0.33</td>
<td>0.17</td>
<td>0.37</td>
<td>0.89</td>
<td>0.39</td>
<td>-0.08</td>
</tr>
<tr>
<td>A/I</td>
<td>0.26</td>
<td>0.24</td>
<td>0.22</td>
<td>0.89</td>
<td>0.39</td>
<td>0.04</td>
</tr>
<tr>
<td>I/A</td>
<td>0.06</td>
<td>0.44</td>
<td>0.02</td>
<td>0.10</td>
<td>0.40</td>
<td>-0.13</td>
</tr>
<tr>
<td>I/A</td>
<td>0.06</td>
<td>0.44</td>
<td>-0.18</td>
<td>0.06</td>
<td>0.44</td>
<td>-0.41</td>
</tr>
<tr>
<td>I/A</td>
<td>0.04</td>
<td>0.46</td>
<td>-0.09</td>
<td>0.07</td>
<td>0.43</td>
<td>-0.03</td>
</tr>
<tr>
<td>NNV: A/A</td>
<td>0.09</td>
<td>0.41</td>
<td>0.33</td>
<td>0.72</td>
<td>0.22</td>
<td>0.30</td>
</tr>
<tr>
<td>A/A</td>
<td>0.10</td>
<td>0.40</td>
<td>-0.04</td>
<td>0.54</td>
<td>0.04</td>
<td>0.34</td>
</tr>
<tr>
<td>A/A</td>
<td>0.10</td>
<td>0.40</td>
<td>0.28</td>
<td>0.78</td>
<td>0.28</td>
<td>0.32</td>
</tr>
<tr>
<td>A/I</td>
<td>0.28</td>
<td>0.22</td>
<td>0.35</td>
<td>0.98</td>
<td>0.48</td>
<td>-0.28</td>
</tr>
<tr>
<td>A/I</td>
<td>0.22</td>
<td>0.28</td>
<td>0.31</td>
<td>0.94</td>
<td>0.44</td>
<td>-0.16</td>
</tr>
<tr>
<td>A/I</td>
<td>0.34</td>
<td>0.16</td>
<td>0.45</td>
<td>0.96</td>
<td>0.46</td>
<td>-0.35</td>
</tr>
<tr>
<td>I/A</td>
<td>0.07</td>
<td>0.43</td>
<td>0.08</td>
<td>0.14</td>
<td>0.36</td>
<td>0.00</td>
</tr>
<tr>
<td>I/A</td>
<td>0.09</td>
<td>0.41</td>
<td>-0.11</td>
<td>0.09</td>
<td>0.41</td>
<td>-0.05</td>
</tr>
<tr>
<td>I/A</td>
<td>0.04</td>
<td>0.46</td>
<td>-0.05</td>
<td>0.21</td>
<td>0.29</td>
<td>0.07</td>
</tr>
</tbody>
</table>

$R_{xy}$ of consistency to latency $= -0.61$.  $R_{xy}$ of consistency to latency $= -0.79$.

Italians have been shown to rely on their second-noun strategy.  In this light, it is interesting that Italians are much faster.  'Good' NNVs produce latencies of -0.35 in Italian versus -0.11 in English; 'Good' VNNs result in latencies of -0.41 in Italian and -0.18 in English.  This apparent advantage holds up only under ideal stress conditions.  Hence it is not animacy alone, but the animacy/stress combination that permits Italians to make faster decisions.
**Competition and conspiracy: effects of stress on worst cases**

In English, the worst cases at the order/animacy level were the competition items in which order and animacy are in conflict: IVA, VAI and AIV. In Italian, the worst cases were generally those in which no animacy information was available at all, i.e., reversible items. How can stress information conspire to make things 'even worse'? Table 6 summarizes the 'best' and 'worst' cases for each word order type in the two respective languages.

<table>
<thead>
<tr>
<th>Language</th>
<th>Word order</th>
<th>Best case</th>
<th>Choice</th>
<th>Latency</th>
<th>Worst case</th>
<th>Choice</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italian</td>
<td>NVN</td>
<td>AVI</td>
<td>+0.49</td>
<td>−0.56</td>
<td>AIV</td>
<td>+0.30</td>
<td>+0.10</td>
</tr>
<tr>
<td>English</td>
<td>NVN</td>
<td>AVI</td>
<td>+0.49</td>
<td>−0.65</td>
<td>VAI</td>
<td>+0.11</td>
<td>−0.06</td>
</tr>
<tr>
<td>Italian</td>
<td>VNN</td>
<td>VIA</td>
<td>−0.44</td>
<td>−0.41</td>
<td>VAA</td>
<td>+0.02</td>
<td>+0.84</td>
</tr>
<tr>
<td>English</td>
<td>VNN</td>
<td>VIA</td>
<td>−0.44</td>
<td>−0.18</td>
<td>VAI</td>
<td>−0.14</td>
<td>+0.51</td>
</tr>
<tr>
<td>Italian</td>
<td>NNV</td>
<td>AAV</td>
<td>+0.46</td>
<td>−0.35</td>
<td>AAV</td>
<td>+0.04</td>
<td>+0.34</td>
</tr>
<tr>
<td>English</td>
<td>NNV</td>
<td>AIV</td>
<td>+0.41</td>
<td>−0.11</td>
<td>AIV</td>
<td>+0.16</td>
<td>+0.45</td>
</tr>
</tbody>
</table>

Starting with the standard order, the worst possible NVN in English is a competition item with stress on the first position: IVA. This is an example of what we called 'a conspiracy of weak vectors': even though SVO interpretation still wins out (82%), the combination of animacy plus contrastive stress on the inanimate object yields the strongest possible competing case for an OVS. Presumably the drop in reaction time (from −0.065 on the best case to 0.00) reflects the time taken to consider this alternative.

In Italian NVN, latency and consistency scores yield different 'worst cases'. Using latency as a criterion, the worst NVN is a reversible item with stress on the second noun: AVA. In the absence of an animacy contrast, this comprises a competition solely between SVO and OVS. *Default* stress would improve the case for SVO, while *contrastive* stress makes OVS more likely. Reaction times must reflect the time taken to compare these two alternatives. However, if we use consistency as a criterion, the worst case is unstressed IVA. Animacy (and hence OVS) wins out in this combination as
well, but only at a rate of 61%. The competing SVO interpretation is apparently helped by an alliance with default stress. Regardless of which criterion we use to determine how stress ‘damages’ reaction times in Italian, one rather general conclusion is clear: unlike their American counterparts, Italians make considerable use of stress information to decide between SVO and OVS alternatives. This is further support for the earlier suggestion that the surface role of subject is more closely associated with topic in Italian, and with agent in English.

Turning to VNN, the worst items in both languages are unstressed: VAI in English and VAA in Italian. Apparently both groups find these orders ‘better’ or easier to resolve when some kind of contrastive stress is available. In fact, the Italian VAA produced the longest reaction times in the whole experiment: +0.84. With stress on the first noun, resolution was faster, with a 63% choice of the second noun; with stress on the second noun, even though reaction times were faster than with no stress at all, performance was still random. Italians are clearly at a standstill on VNN items, caught between two equally plausible VOS and VSO alternatives. In fact, there is actually no evidence for a preferred word order in Italian VNN. The only bias is toward a particular order/stress combination: VOS. Without stress, the order information is uninterpretable.

On the NNV items, the English worst case is AV and the Italian worst case is AAV. The American case is another example of stress shoring up the losing alternative in a competition, allying with animacy to increase the probability of SOV over the prevailing OSV tendency. The Italian case is another example of Italian difficulty with reversible items. In this particular word order Italians do have a preference for SOV. However, this preference is completely wiped out when contrastive stress falls on the first noun, shoring up the likelihood of an OSV interpretation.

To summarize the effects of stress in this experiment, Italians clearly make much greater use of the pragmatic information in contrastive stress to interpret sentences out of context. Stress can cancel out slight word order biases, and significantly reduce the effects of a standard SVO interpretation. For Americans, decisions are affected relatively little by the presence or absence of contrastive stress. However, reaction time data show that the competing and converging information in stress does register for the Americans as well. Finally, we can make no strong conclusions about the direction of these stress effects. In general, stress can retard or accelerate decisions by allying with the winner or the loser of an order/animacy standoff. Default stress accelerates the interpretation of standard orders while marked stress seems to help in interpreting non-standard orders. Beyond these statements, more specific conclusions (i.e., whether object stress is invariably better than
subject stress) must await further study. However, we definitely feel that the competition/convergence analysis is a useful tool in understanding such complex effects.

Correlations between latency and choice

We have argued that both the outcome and the time taken to reach a decision should be a function of the number of plausible alternatives, and the relative weights of converging and competing vectors. This leads us to predict a significant negative correlation between response latency and response consistency, at each level of the analysis. Slow reaction times and low consistency scores reflect the plausibility of competing alternatives.

At the three-way level (language $\times$ animacy $\times$ order) there were nine possible item types (summing across stress and topic conditions) within each language. Pearson product-moment correlations were calculated for each language group, comparing latency and response consistency for each of the nine item types. The results were correlations of $-0.82$ for English and $-0.92$ for Italian (both significant at $p < 0.01$). In other words, the choice and reaction time data are compatible.

At the four-way level (language $\times$ animacy $\times$ order $\times$ stress), there were twenty-seven possible item types in each language (summing across topic conditions). Pearson product-moment correlations were calculated once again between consistency and reaction time means. The resulting correlations were $-0.61$ in English and $-0.79$ in Italian, both significant at $p < 0.01$. Hence even at the four-way level, the two measures yield similar results—but to a lesser degree once stress enters the picture. For one thing, the effect of stress was so small (particularly in English) that the spread between mean choice scores as a function of stress is often negligible, and perhaps unreliable. Nevertheless, the results are compatible with the interpretation that reaction times reflect the weighing of alternatives.

Induced introspection in relation to behavior

The induced introspection procedure was included for two reasons, one methodological and the other theoretical. On methodological grounds, we wanted to counter the possibility that apparent language differences were really artifacts of cultural differences in interpreting the goal of the experiment. If we listed explicitly all the factors that were relevant to sentence interpretation, then presumably all the subjects would know that these
factors were fair game. In light of the dramatic language differences we obtained in the use of animacy versus word order, this was clearly an important control. On a more theoretical level, these judgments give us interesting information about metalinguistic processing, i.e., how well subjects can reflect on their own processing strategies in a sentence comprehension experiment. Finally, we can compare our own inferences about behavior in this experiment with the subjects' impressions of how competing and converging sources of information affected their decisions.

Ideally, we would have liked to obtain a rank-order judgment for each of the 81 sentence types. This was clearly impractical, so we compromised by dividing items into blocks of nine, each block containing all nine order/stress combinations in one animacy condition and one topic condition. At the end of each block, subjects ranked the 3 to 5 factors that had varied across items: order, stress, different kinds of verbs, and (if relevant to that block) topic and animacy. Because the range of scores varied from 3 to 5, all scores were prorated to reflect a hypothetical 0–5 scale. Since tie ranks were permitted, the rank orders for each factor were in principle not mutually exclusive.

Table 7 summarizes three kinds of information obtained with these rank orders: mean ranks for each factor (across nine blocks), internal consistency

| Table 7. Induced introspection judgments: Internal consistency and correlations with behavior |
|-----------------------------------------------|--------------|
|                                                | English      | Italian      |
|                                                | Mean rank orders across nine blocks | Mean rank orders across nine blocks |
| Order  | Animacy | Verbs   | Stress | Topic | Order  | Animacy | Verbs   | Stress | Topic |
| 3.42   | 2.20    | 2.30    | 1.82   | 1.11   | 2.23   | 3.43    | 2.72    | 2.31   | 1.74  |
| 0.54*  | 0.16    | 0.54*   | 0.49*  | 0.77*  | 0.42*  | 0.44*   | 0.55*   | 0.44*  | 0.28  |

Correlations of each judgment mean rank and corresponding choice behavior

| Order  | Animacy | Verbs   | Stress | Topic |
| 0.54*  | 0.48*   | –       | 0.25   | 0.51* |
|        |         |         | 0.42*  | 0.78* |

*p < 0.05.

of ranks for each factor (i.e., correlation of each block with the mean ranks using an item-scaling analysis), and correlations of mean rank orders for each factor with corresponding choice behavior in the experiment.

First of all, the mean rank orders were consistent with behavioral results in the analyses of variance. In English, the subjective ratings from highest to
lowest were as follows: order, verbs, animacy, stress and topic. In Italian, the corresponding ranks were: animacy, verbs, stress, order, topic. In other words, English listeners knew that they were relying primarily on word order; Italians knew that they placed greatest weight on lexical information. And both groups knew that the topic manipulation had little effect on them. The only discrepancy between the introspection data and the analyses of variance is that Italians overestimate the weight they place on stress and underestimate the effects of order.

The internal consistency of rank orders among the nine blocks was significant, but not overwhelming. In other words, subjects changed their minds about their ‘preferred’ factors from one block to another. This is, in part, to be expected, since the blocks varied in whether or not topic and animacy information were available. Also, the variability from one block to another means that subjects were continually weighing information; they did not choose one single strategy and stick to it obstinately throughout the experiment.

Finally, the mean rank for each factor, for each subject, was compared with his use of that factor in the experiment. To calculate this relationship, subjects received scores for the number of times they chose the animate element, the element corresponding to word order preferences in his group, the stressed element, and the topicalized element. As can be seen in Table 7, subjects do know to some extent what they are doing. The correlation between order ranks and order-based choice was 0.54 in English and 0.48 in Italian. Corresponding ranks for animacy were 0.48 in English and 0.78 in Italian. At the level of topic and stress, this pattern breaks down, so that mean rank orders and choice behavior are unrelated. However, the low correlations for topic and stress fit with the interactions in the analyses of variance, since both stress and topic were secondary vectors that had effects only on certain combinations of order and animacy. Hence, ranked across blocks of nine items, these judgments could not reflect choice behavior.

In sum, the induced introspection procedure works. When their attention is explicitly directed to the factors that converge and compete in this experiment, adult subjects are able to reflect accurately on their own behavior—at least with regard to the two primary vectors of word order and animacy. These segments also brought out some interesting anecdotal information. Italians occasionally applied their animacy bias in a very broad and ‘blind’ way. One actually told us, with some irritation, ‘Look, if you give me any more pencil/cow items, I can tell you right now how they’ll come out’. However, even the most avid users of lexical biases stopped short on certain contrasts, e.g., ‘The cube hit the cow’. That is, certain verbs do permit a plausible interpretation with the inanimate objects as the actor. Since par-
ticular noun/verb combinations were assigned in individual random combinations for each subject, such plausibility differences could not be reflected in the experimental analyses. However, the anecdotal reports suggest that the plausibility issue is worth pursuing in further studies. Plausibility apparently does not correspond to a 'crisp' animate/inanimate distinction. A particularly interesting report in this regard came from an Italian subject who was faced with the item ‘The cube grabs the monkey’. The prevailing tendency among Italians in general, including this woman, would have been to select the monkey. In this case, she chose the cube, in keeping with an SVO interpretation. During the induced introspection segment that followed she reported the following reasoning:

'The movie King-Kong? Well, if you think of the skyscraper as a cube, with that big monkey caught up on top of it, then you could say “The cube grabs the monkey”.

With regard to the American anecdotal reports, it appears that these subjects knew they were applying word order but did not exactly know how or why. SVO word orders obviously fit a layman’s view of ‘good English’. But the second-noun strategy, as far as we know, has never been reported in the linguistic or psycholinguistic literature. It is unlikely, then, that Americans could know why they were behaving so systematically. None of them volunteered a characterization of this strategy.

Performance without induced introspection

The success of the induced introspection procedure leaves us with another problem. Perhaps this procedure is so powerful that it induces artificial comprehension strategies that would never have occurred in more traditional sentence comprehension experiments. To control for this possibility, and explore the generalizability of the findings, we conducted a second experiment in English without the induced introspection procedure. The experiment was altered slightly: all items were randomized (as opposed to the block format used in Experiment 1), and we tried a somewhat different version of the topic manipulation. We did not list the factors to be varied in advance, nor ask for judgments after items were administered.

Subjects were, once again, 30 undergraduates at the University of Colorado, all native speakers of English. The same within-subjects analyses of variance, on choice and latency, were conducted. In the interests of brevity, we will not present details from that experiment here. The important point for present purposes is that the results of Experiment I were replicated.
There were significant main effects for all the variables, although stress and the slightly altered topic manipulation continued to be secondary determiners of sentence interpretation. In this second sample of English listeners, word order again played a much larger role than animacy. Most important, the second-noun strategy uncovered in Experiment I appeared again in this experiment, just as systematically. Hence the VOS/OSV strategies are not artifacts of telling subjects to pay attention to order. English listeners reinvent it regardless of instructions.

Since cross-linguistic experiments are costly and time-consuming, we did not carry out the second experiment in Italian as well. But we did carry out further pilot testing of a variety of instructional conditions with a separate sample of Italian listeners, without induced introspection. Regardless of instructions, Italians continued to rely overwhelmingly on the animacy contrast. Nor did we find evidence for non-canonical strategies that might have been missed in Experiment I. This was true even for a group of linguists who submitted to a shortened version of the test. We concluded that the major effects of Experiment I are in no way artifacts of instructional conditions.

We also tried to explore the relationship between use of semantic strategies and judgments of grammaticality in Italian, by asking a separate group of nine subjects for ratings of the experimental sentences on (1) plausibility and comprehensibility, versus (2) ‘Is this an acceptable, grammatical sentence in Italian?’ The overlap between the two ratings was almost 100%. Educated Italian laymen apparently equate grammaticality and lexical plausibility when word order and pragmatic factors are varied.

Individual differences within languages

Gleitman and Gleitman (1970) report a wide range of individual differences within the English-speaking population in the degree to which semantic factors (i.e., plausibility of certain lexical combinations) influence decisions. From this point of view, Italians as a group could be compared with that subgroup of English speakers who rely most heavily on lexical factors. Similarly, Americans as a group may look like the extremes of the Italian population in their reliance on word order. If we could locate the ‘outriders’ in our two populations, the two groups might end up looking as though they spoke the same type of language. English and Italian speakers do, as we have shown, share the same prototype: an SVO order, with unmarked stress, in which the first noun is animate and the second inanimate. It is possible, for example, that the second-noun strategy which dominated American perfor-
mance is a by-product of strong reliance on this SVO prototype. That is, VOS and OSV biases might not have an independent psychological status of their own; instead, they may be derived on-the-spot from this SVO base, taking the verb as an anchor point and looking outward for the nearest noun. Hence the preverbal noun in an NNV would have to be the subject, and the immediate post-verbal noun in a VNN would have to be the object. If this is the way that the second noun strategy is derived, then we might expect Italians at the extreme end of reliance on SVO to show a similar strategy.

Another issue has to do with the relative use of contrastive stress in the two language groups. Since Italians demonstrably rely more on semantic-pragmatic information, their greater use of stress may reflect a search for pragmatic interpretations. If we could find those Americans who make greatest use of lexical information, they may also rely on stress to a larger degree, comparable to their Italian counterparts.

The basic question at issue is this: are the various competition/convergence patterns described so far a product of individual differences in the use of semantic information? Or do they reflect true differences in the psychological underpinnings of English and Italian?

To explore this question, we located the ten ‘least typical’ members of each of the two language groups. The ‘least English’ were those who made least consistent use of the dominant bias toward word order (including first noun choice in NVN, and second noun choice in the other two orders). The ‘least Italian’ subjects were those who made the least consistent use of the dominant bias toward animacy. With these two groups, we repeated the $2 \times 3 \times 3 \times 3$ (language x word order x animacy x stress x topic) analysis of variance on percent choice first noun. We did not repeat the latency analysis on this subgroup; because the original latency analysis accounted for much less variance overall, it seemed imprudent to repeat the analysis with decreased power in a comparison of two 10-subject groups.

The results of the analysis of variance are summarized in Table 8; cell means for the four-way language x word order x animacy x stress interaction are presented in Table 9. Most of the results duplicated those for the larger sample. All five main effects reached significance, in the direction described earlier. At the 2-way level, the only change was that the language x animacy interaction dropped out. This means that we did indeed succeed in equating the two subgroups for their relative reliance on semantic information. However, the other two-way interactions were exactly as reported before. In particular, stress continued to have a greater impact in Italian while order continued to have a larger effect in English.

At the three-way level, the language x word order x topic interaction also dropped out. Hence we seem to have equated the subgroups not only for use
of animacy, but for reliance on topic information. However, the other three-way interactions remained: language \times \text{word order} \times \text{animacy}, and language \times \text{word order} \times \text{stress}. These two in turn entered into a four-way interaction: language \times \text{word order} \times \text{animacy} \times \text{stress}. Results were equivalent to those presented for the full sample.

In other words, the structural differences between Italian and English did not disappear when we artificially equated subjects for their use of semantic information. In their relative use of SVO, the two subgroups look alike. In fact, the Italians 'out-English' the English, with word order winning over animacy in two of the three IVA competition items: unstressed IVA and final-stressed IVA. But when stress falls on the inanimate object, in IVA, we have an alliance between animacy and object stress that defeats SVO in Italian but not in English. No matter how much we 'squeeze' these subgroups together, basic word order is still subject to pragmatic influences in Italian.

Finally, the least English subjects continue to use a second-noun strategy that does \textit{not} appear in the least Italian group. In the English subgroup, VOS and OSV lose out to animacy in competition items (in contrast to the larger sample). But the size of the animacy victory is much smaller than the corresponding animacy effects with the Italian subgroup. The least Italians still have the \textit{opposite} bias on NNV compared with Americans: an SOV preference that puts up appreciable competition against animacy. And on VNN items, they continue (like the larger sample of Italians) to have no obvious word order preferences in the absence of stress.

### Table 8. Summary of significant results from the ANOVA on percent first noun choice in the 10 'least Italian' and 10 'least English' subjects

<table>
<thead>
<tr>
<th>Main effects &amp; interactions</th>
<th>d.f.</th>
<th>F-ratio</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>1,18</td>
<td>16.78</td>
<td>0.001</td>
</tr>
<tr>
<td>Word Order</td>
<td>2,36</td>
<td>55.03</td>
<td>0.000001</td>
</tr>
<tr>
<td>Animacy</td>
<td>2,36</td>
<td>57.82</td>
<td>0.000001</td>
</tr>
<tr>
<td>Topic</td>
<td>2,36</td>
<td>7.67</td>
<td>0.003</td>
</tr>
<tr>
<td>Language \times \text{Word Order}</td>
<td>2,36</td>
<td>6.32</td>
<td>0.005</td>
</tr>
<tr>
<td>Language \times \text{Stress}</td>
<td>2,36</td>
<td>2.73</td>
<td>0.03</td>
</tr>
<tr>
<td>Word Order \times \text{Animacy}</td>
<td>4,72</td>
<td>9.90</td>
<td>0.00003</td>
</tr>
<tr>
<td>Word Order \times \text{Stress}</td>
<td>4,72</td>
<td>5.21</td>
<td>0.002</td>
</tr>
<tr>
<td>Word Order \times \text{Topic}</td>
<td>4,72</td>
<td>2.57</td>
<td>0.05</td>
</tr>
<tr>
<td>Language \times \text{Word Order} \times \text{Animacy}</td>
<td>4,72</td>
<td>3.87</td>
<td>0.007</td>
</tr>
<tr>
<td>Language \times \text{Animacy} \times \text{Stress}</td>
<td>4,72</td>
<td>4.34</td>
<td>0.004</td>
</tr>
<tr>
<td>Language \times \text{Word Order} \times \text{Animacy} \times \text{Stress}</td>
<td>8,144</td>
<td>2.16</td>
<td>0.04</td>
</tr>
</tbody>
</table>
In conclusion, the ‘functional map’ that we have drawn of competing and converging vectors in the two different languages seems to have a validity that transcends individual differences in the weights of those vectors. Stress has more effect on the English subgroup than it has on the larger sample of English speakers; but stress continues to have a larger effect in Italian. Both groups, when equated for use of SVO and animacy, continue to use different strategies to interpret non-canonical orders. If the second-noun strategy is a spur-of-the-moment invention by English listeners, operating on a strong SVO prototype, then it is not obvious why Italians with
an equally strong SVO prototype do not use the same strategy to derive their interpretations. We think, instead, that the word order strategies displayed by the two different language groups have something to do with the alternative pragmatic uses of word order that occur in spontaneous, informal production. Where those alternatives come from is another matter.

Summary and general discussion

The main findings of this study can be summarized as follows.

1. Word order and animacy are the major factors determining sentence interpretation in both Italian and English. Topic and stress are 'secondary vectors' that operate primarily by allying with animacy, order, or both.

2. However, there is a dramatic 'mirror image' difference between the two languages in the degree to which they depend on the information in word order versus the information in lexical configurations. Americans depend far more on order; Italians depend far more on semantic factors. Furthermore, Italians also make greater use of the secondary information in topicalization and stress, suggesting that the notion of 'subject' is more closely tied to topic in this language.

3. The English reliance on word order extends not only to SVO, but to a very consistent second-noun strategy with non-canonical orders, yielding VOS and OSV interpretations. No such biases are found in Italian in VNN, and they have a slight bias in the opposite direction from Americans in NNV (yielding SOV interpretations). These non-canonical patterns mirror the range of word order variations that are found in both languages, in informal speech production. English can permit VOS and OSV (right-dislocation and left-dislocation) but not VSO and SOV; Italian permits all possible word order combinations, including OVS, under certain semantic-pragmatic conditions.

4. The effects of contrastive stress are not unidirectional. Rather, default stress produces better performance in NVN items and marked stress produces better performance on alternative orders. However, these general tendencies also depend on the kinds of alliances being formed with animacy versus word order in all these sentence types.

5. Reaction time and response consistency data support a competition/convergence analysis: the more sources of information converging on a single interpretation, the faster and clearer the response; weak vectors take longer to resolve than strong ones in the same direction; and even though the outcome of a decision process may ultimately go with the strong 'winning'
vector, conspiracies of weaker vectors can shore up the probabilities of alternative interpretations and hence slow down reaction times.

(6) Results of the induced introspection procedure show that both Americans and Italians are aware of their respective hierarchies of strategies. Anecdotal information from those segments also suggests that Italians are not only relying on a crisp animate/inanimate division, but also calculate the relative probabilities of specific lexical combinations (e.g., the King Kong example).

(7) A control experiment and some subsidiary analyses showed that the above results are not artifacts of instructions. And even when individual differences in degree of semantic versus SVO bias are taken into account, the direction of the language-specific results (i.e., use of stress and non-canonical strategies) remains the same.

The results presented here are compatible with a functionalist theory that accounts for differential weightings of grammatical cues in a principled way. As we understand them, deterministic theories would make no predictions concerning the relative strength of converging cues, nor the outcomes of competitions and conspiracies among cues. To preserve a competence model which characterizes grammatical knowledge as a set of discrete, deterministic rules, these findings would have to be relegated to a separate performance component. Until the shape of that component is specified, the deterministic model cannot be tested or evaluated against these or any other data.

In addition, our cross-linguistic results pose problems for the universality of models that require a strict separation of semantic, pragmatic and syntactic processes. Consider Gleitman and Gleitman's proposal concerning a distinction between core and peripheral rules. It does seem fair to conclude that word order stands at the 'core' of English, insofar as NVN order produced fast reaction times and consistent response even in the face of conflicting information from lexical items. Certainly semantics can contribute to the interpretive process, speeding or slowing reaction times. But word order always 'wins' in English. However, if we use consistency and latency data as a criterion for assignment to the core or the periphery, then we must also conclude that semantic plausibility stands at the core of Italian, with word order employed as a secondary or peripheral strategy. These conclusions apply not only to competition situations, which are ambiguous by definition, but to convergence items where there is no conflict at all between semantic and syntactic information. Even on the reversible AVA items, Italians were slower and less consistent in applying SVO interpretations than Americans. Clearly SVO is not at the core of Italian sentence processing.
This brings us to another boundary issue, regarding the status of non-canonical word order strategies in the grammar. We have suggested that the English second-noun strategy is related to the occurrence of left- and right-dislocations (OSV and VOS) in informal speech production. Similarly, the fact that Italians were unable to resolve non-canonicals without additional cues is related to the fact that all word orders are possible in informal Italian speech production. If the alternative word orders have systematic effects in both comprehension and production, then we submit that such grammatical knowledge belongs in the grammar. On the other hand, some of the models discussed earlier would presumably place these non-canonical strategies into some kind of a *semi-grammar*. The only allusion we have found in the literature to a non-canonical sentence strategy in English adults is Forster's prediction cited earlier that listeners will use a first noun strategy to deal with ungrammatical sentences like 'Doctor patient cure'. This is in fact the *opposite* of what we found. Furthermore, the second-noun strategy was extremely robust (averaging 85% in two experiments in English). If we decide to put the English OSV and VOS tendencies into a semi-grammar, then the contrast between English and Italian presents a serious problem. English use of OSV and VOS was more consistent, and almost as rapid, as Italian use of SVÒ. By what criterion do we classify English OSV and VOS as *semi-grammatical* while at the same time retaining SVO within the *grammar* of Italian? We might end up having to classify all Italian word order phenomena as penumbral, peripheral, or semi-grammatical.

However, suppose we do decide that non-canonical strategies belong in the grammar? Antinucci (1977) and Duranti and Ochs (1979), have shown that at least some of the systematic order variation in Italian can be described if pragmatic distinctions like given and new are included as central mechanisms in the grammar. The same kinds of constraints seem to operate in the production of left- and right-dislocations in English. This is, in fact, the solution that we suggest: building systematic facts about semantic and pragmatic contributions to syntactic processing into one unified model of grammar.

One still might want to argue that we have found a *cultural* difference, not a linguistic one. Are Italians merely 'soft on syntax'? Is Mediterranean 'gioia di vita', a common sense approach, pitted against Anglo-Saxon orderliness and syntactic conservatism? First of all, if this were true, we still could not explain why Italians had difficulty applying SVO in reversible sentences where common sense didn't get in the way at all. Second, perhaps it has been anglo-centric of us to have identified 'syntax' with word order. In the English-based psycholinguistic literature, 'semantic' or 'probable event' strategies have typically been identified with the generic outgroup: with
children below the age of 5 (e.g., Strohner and Nelson, 1974), with aphasic adults (Zurif, Caramazza and Myerson, 1972), and in Gleitman and Gleitman's nominal compound task, with high-school educated adults compared with graduate students. If we follow this historical line into the present study, we are in the embarrassing position of putting educated middle-class citizens of a major Western developed nation in the generic outgroup. What if we were to give Italians a piece of more reliable grammar to work with? Recall that all of the 81 items in the experiment in both languages, were ambiguous with regard to subject-verb agreement. We are currently conducting an experiment that sets order, animacy and stress information in convergence/competition relationships with verb agreement, e.g.,

(19) Is eating the horses the cow. Mangia i cavalli la vacca.
(20) The cube are grabbing the monkeys. Il cubo afferrano le scimmie.

Preliminary results suggest that English and American subjects react quite differently. Americans apparently do not feel bound to respect agreement when order stands in opposition. However, agreement seems to be more deterministic in Italian. We suggested earlier that Italian is 'less' of an SVO language than English. It may also be 'more' of an inflectional language than English. There may be a continuum between word order and case languages, with Italian located somewhere in between.

We are now carrying out some production and acceptability judgment studies using the same competition/convergence approach adopted here, in the hope that results from different performance domains can be unified within a single, coherent performance grammar (or at least a fragment of a grammar for certain topic/agent/subject phenomena). This must include an exploration of some new mathematical models to represent and test the interacting vector weights that have been postulated to describe probabilistic form/function mappings (e.g., MacWhinney and Connell, in preparation). Without denying how much more work has to be done before anything like a reasonable performance grammar is available, we hope at least to have underscored the value of cross-linguistic comparisons in studies of sentence processing, in testing the universality of any processing theory.

References


References Notes


L'anglais et l'italien présentent des contrastes intéressants et pertinents pour un problème crucial en psycholinguistique, celui de la frontière entre connaissance grammaticale et extragrammaticale dans le traitement des phrases. Bien que tous deux soient des langues avec un ordre SVO sans inflexions de cas pour indiquer les relations grammaticales de base, l'italien autorise beaucoup plus de variations dans l'ordre des mots pour des buts pragmatiques. Les italiens doivent, donc, s'appuyer plus que les anglais sur des facteurs autres que l'ordre des mots. Dans l'expérience présentée, on a demandé à des adultes anglais et italiens d'interpréter 81 phrases simples où variaient l'ordre des mots, les contrastes animés/non animés entre deux noms, le stress contrastif et la topicalisation. Les italiens s'appuient principalement sur des stratégies sémantiques alors que les auditeurs anglais s'appuient sur l'ordre des mots et cela inclue une tendance à interpréter le second nom comme sujet dans les ordres de mots non-canoniques (correspondant aux variations d'ordre de la production de l'anglais informel). Les italiens font un plus grand usage du thème et de l'information donnée par l'accent. Enfin, les italiens sont beaucoup plus lenis et moins consistants dans l'application de stratégies d'ordre de mot même pour des phrases reversibles NVN où il n'existe pas de conflit entre l'ordre et la sémantique. Cela suggère que l'italien est 'moins' une langue SVO que l'anglais. Les stratégies sémantiques tiennent apparemment au 'coeur' de l'italien les mêmes rôles que les stratégies d'ordre des mots au 'coeur' de l'anglais. Ces résultats font problème pour parler d'une séparation 'universelle' entre sémantique et syntaxe et pour les théories qui postulent une priorité 'universelle' d'un type d'information sur l'autre. Les résultats sont examinés dans le cadre d'un modèle de compétition, approche fonctionnaliste de la grammaire qui rend compte de façon rigoureuse des données probabilistes et des poids différents des différentes sources (converges et rivales) d'information dans le traitement des phrases.