LANGUAGE AND COGNITIVE PROCESSES, 1999, *14* (2), 129–171

Processing of Morphological and Semantic Cues in Russian and German

Vera Kempe
State University of New York at Oswego

Brian MacWhinney
Carnegie Mellon University

This study examines the on-line processing of morphological cues to sentence interpretation in Russian and German with the goal of evaluating the relative impacts of cue availability and cue reliability. Both Russian and German use the cues of word order, animacy, case-marking, and subject–verb agreement to identify the agent of active transitive sentences. However, the availability of the case-marking cue is higher in Russian than in German. Using a picture-choice paradigm, we contrasted case-marking and animacy in Russian and German. The reaction times showed larger effects of case-marking in Russian than in German and effects of animacy in German, but not in Russian. These results suggest that the higher the availability of a cue, the larger the processing benefits associated with the presence of this cue and the smaller the impact of other converging information. A recurrent cascaded backpropagation network was designed to simulate these effects. The network succeeded in capturing the essential language differences in the reaction times, thereby illustrating how the statistical properties of cues in a language can affect the time-course of activation of alternative interpretations during sentence processing.

INTRODUCTION

Crosslinguistic studies of sentence processing have documented pervasive differences between languages (Cuetos & Mitchell, 1988; Frazier & d'Arcais, 1989; Vigliocco, Butterworth, & Semenza, 1994). The Competi-

Requests for reprints should be addressed to Vera Kempe, Department of Psychology, SUNY Oswego, Oswego, NY 13126, USA. E-mail: kempe@oswego.edu

This research was supported by a fellowship from the German Academic Exchange Board (DAAD) to the first author. We would like to thank Don Mitchell, Roman Taraban, and five anonymous reviewers for helpful comments on an earlier draft of this paper.

tion Model (MacWhinney & Bates, 1989) has been proposed as a way of relating the observed processing differences to variations in language structure. In particular, the Competition Model focuses on the role of surface cues such as word order, noun animacy, subject–verb agreement, and noun case-marking. The strength of a cue is viewed as depending on three factors: (i) its availability, defined as the proportion of times a cue is present and can be used for accessing the underlying function, (ii) its reliability, defined as the proportion of times a cue signals the correct interpretation given that it was present, and (iii) its cost, which depends on the perceptual salience of the cue and the load it places on working memory.

Studies within the Competition Model framework have examined processing in over 15 languages (Bates, McNew, MacWhinney, Devescovi, & Smith, 1982; Kail, 1989; MacWhinney, Bates, & Kliegl, 1984; MacWhinney, Pléh, & Bates, 1985; McDonald, 1986; Sokolov, 1989). In a typical Competition Model experiment, subjects are presented with simple transitive sentences and are asked to decide which noun refers to the agent of the sentence. A consistent finding from these studies is that, if the cues for agentivity are placed in competition with each other, the choice of one of the nouns as the agent can be modelled by multiplicative cue integration (McDonald & MacWhinney, 1989). In these off-line studies, it has been found that the primary determinant of cue strength is cue reliability. In other words, when subjects are given enough time to permit full, deliberate consideration of all competing cues and interpretations, they integrate cues in a way that maximises the probable correctness of their final interpretation.

Newer work in the Competition Model framework (Hernandez, Bates, & Avila, 1994; Kilborn, 1989; Li, Bates, & MacWhinney, 1993; McDonald & MacWhinney, 1995; Mimica, Sullivan, & Smith, 1994) has attempted to extend the earlier off-line models to the study of on-line sentence processing. The paradigm used in many of these studies is a speeded cross-modal picture-choice task where participants listen to a transitive sentence while making a choice between two candidate agents that are presented visually. Both candidates are named in the sentence, but only one can be selected as the agent. Participants are encouraged to respond even before the sentence is finished, if they are confident of their interpretation. The use of this speeded picture-choice task provides insights into how the complex interaction between various cues unfolds over time. There are two ways in which the results from this on-line technique further support the emphasis on cue reliability that emerged from the earlier off-line studies. First, it has been found that the strongest and most reliable cues lead to the fastest reaction times. Second, it has been found that any type of competition or disagreement between cues results in inhibition and slower reaction times. These two effects are well-predicted by a simple multiplicative cue integration model in which cue strength is largely predicted by reliability.

However, there are two other effects found in these on-line studies that point towards the need for a more complicated model. First, certain cues have a more pronounced effect on the reaction times than on the choice of one of the nouns as an agent. Second, reaction times are not as clearly affected by cue convergence as are patterns of agent choice. It appears that strong cues tend to saturate the on-line processing system, so that providing additional evidence when a strong cue is already present has little additional effect on reaction times. There may also be a trade-off between the benefits gained from obtaining more evidence from a convergent cue and the costs associated with processing this cue (Kail, 1989; Mimica et al., 1994).

We noted earlier that the Competition Model relates cue strength to two separate cue validity factors: reliability and availability. Reliability is a measure of cue consistency and dependability, whereas availability is a measure of frequency. There is some reason to believe that the frequency of a syntactic structure has a stronger effect upon on-line role assignment processes than upon off-line interpretation. In related work, frequency effects have been demonstrated for relative clause attachment (Cuetos & Mitchell, 1998; Mitchell & Cuetos, 1991; Mitchell, Cuetos, & Zagar, 1990), as well as for the processing of syntactically ambiguous lexical items under the influence of verb frame structures (MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Garnsey, 1994; Trueswell, 1996). This related literature would lead us to suspect that frequency would also play a major role in determining agent role assignment. In fact, the model developed by MacDonald et al. (1994) corresponds quite closely to one offered by MacWhinney (1987). In both accounts, alternative attachment structures compete in terms of alternative lexical forms or homonyms. The candidacy of each competing attachment is supported by cues that vary in strength. MacDonald et al. (1994) link cue strength more to frequency, whereas the Competition Model has often linked it more to reliability. It would be interesting if this contrast reflects the greater emphasis given to on-line data by MacDonald et al. (1994) or Mitchell et al. (1990). If so, we would expect to find a stronger effect of frequency on Competition Model experiments when they include a stronger on-line focus.

It is possible to disentangle the effects of frequency and reliability by utilising the fact that many languages have fully grammaticised cues that are completely reliable, but not fully available. The specific inflectional paradigm of the language determines the extent to which a fully reliable cue is also available in particular sentences. A comparison of two languages in which reliable morphological cues differ only in availability allows us to measure the impact of availability during on-line sentence processing. In this study, we will compare on-line processing in Russian and German, two languages which have an identical repertoire of cues for agent role assignment, but which differ with respect to the availability of morphological cues. If processing differences between these two languages can be found, this will provide evidence for differential effects of cue availability during on-line sentence processing.

The distribution of cues in Russian and German

The present study focuses on the cues for agentivity in simple transitive sentences that take the form non-verb-noun (NVN). A typical sentence of this type is ‘The boy kicked the ball’. This sentence type was chosen because it is the one used in many crosslinguistic studies that have been conducted in the past.

The cues provided in Russian and German for determining who did what to whom in a simple active transitive sentence are word order, animacy, case-marking and subject–verb agreement. We will first describe each of these four cues and then provide quantitative information about their availability and reliability in Russian and German.

Word Order. The ordering of the nouns and the verb in a sentence can provide information about which noun is most likely to be the agent. It has been shown, for example, that positioning of a noun before the verb is a very strong cue for agentivity in English (MacWhinney et al., 1984). In German, there is a strict rule that requires that the tense-bearing verb should appear in second position. If tense is marked by an auxiliary verb, it is this verb that appears in second position and other verbs will appear in final position. As long as the tense-second rule is obeyed, German permits variation in the placement of the nouns and other arguments around the verb. In particular, both SVO (subject-verb-object) and OVS (object-verb-subject) orders are permitted in German main clauses, although SVO order is much more common than OVS order. The selection of either SVO or OVS is based on pragmatic factors. Because of the verb-second rule, German does not typically permit SOV or OSV in main clauses. However, the order AVSO can arise when an adjunct (adverb, participle, or prepositional phrase) is placed in first position. In such cases, AVSO is strongly preferred over AVOS. Russian, on the other hand permits all six basic word orders (SVO, OVS, VSO, VOS, SOV, OSV) and the selection of any one of these orders is based on pragmatic factors. For the SVO and OVS sentences that will be the focus of the current study, both languages make a choice between the two orders on the basis of pragmatic factors.

Animacy. There is a general tendency in all languages for agents to be animate. In a sentence with two nouns that differ in relative animacy, the animacy cue can be used to favour the selection of the most animate noun as agent. In most sentences, the availability of the animacy cue tends to converge with and blur into the 'probable event' cue, since the most animate noun is usually most likely to be the actor. However, Competition Model studies that have been devoted to disentangling the relative effects of the two cues indicate that animacy is the dominant cue (Bates et al., 1982; Bates, MacWhinney, Caselli, Devescovi, Natale, & Venza, 1984). In the current study, we will be examining a fairly rigid form of the animacy cue, since we will not be including a large range of noun-verb combinations. As Corrigan (1986, 1988) has shown, animacy information interacts strongly with the details of the relations between specific nouns and specific verbs. However, the details of these interactive processes lie outside the scope of the current study.

The availability of animacy contrasts depends on the pragmatics of the discourse context and is unlikely to differ significantly across languages. However, it has been shown that languages differ markedly in terms of the actual reliance that they place on the animacy cue, when it is present. Languages such as Italian, Spanish, and German (e.g. Bates et al., 1982; MacWhinney et al., 1984; Hernandez et al., 1994) place a greater reliance on animacy than languages such as English and Hungarian. In the case of English, it appears that the strong word order cue completely obviates the need to rely on additional information from animacy. Similarly, in Hungarian, the strong case-marking cue makes animacy relatively less important. In Italian, Spanish, and German, on the other hand, the occasional absence of some other highly determinant cue increases the role that the animacy cue can play during sentence processing.

In order to fully evaluate the evidence provided by an animacy contrast, the listener has to process the whole clause. However, in the speeded picture-choice paradigm, the animacy of each individual noun can contribute incremental evidence to the ongoing decision process. Because of this possible incremental use of animacy, our corpus analyses code the reliability of the animacy cue in terms of its effect on the choice of individual nouns as agent.

Case-marking. Unlike a pragmatic cue, such as animacy, the case-marking cue is a fully conventionalised cue that can independently signal the case role status of each noun phrase. Although this cue is fully grammaticised, it may often be ambiguous or absent. In German, case-marking varies according to the number and gender of the noun. The actual markings are mainly expressed on the articles or adjectives that precede the noun. For some cases and some nouns, case is also marked on suffixes. In Russian, case-marking varies according to the number, gender, and animacy of the noun. The forms used to mark case appear exclusively as suffixes that follow the noun. Although both languages possess nominative and accusative marking, there are significant differences in the patterns of neutralisation within the two declensional paradigms. These differences are crucial to the current study, because they account for the differential availability of nominative and accusative marking in Russian and German. In German, nominative and accusative are neutralised in feminine (*die-die*), neuter (*das-das*), and plural (*die-die*) nouns. Thus, in order to reliably indicate either nominative (*der*) or accusative (*den*) marking, at least one masculine noun has to be present in the sentence. In Russian, nominative-accusative neutralisation occurs mainly when there are masculine inanimate or neuter nouns which end in a consonant. In these nouns, genitive, dative, instrumental, and locative are marked by different vowel suffixes, whereas nominative and accusative both end in a null-morpheme. Additionally, feminine nouns ending in an end-palatalised consonant (e.g. *mat'*),[1] also exhibit nominative-accusative neutralisation. Since end-palatalised feminine nouns are quite infrequent, and most Russian neuter nouns are inanimate, nominative-accusative neutralisation mainly co-occurs when both nouns are inanimate. However, text counts (Zubin, 1977, 1979) show that sentences with two inanimate nouns and a transitive verb are rather rare in languages. These different neutralisation patterns and the ways in which they interact with animacy configurations are responsible for the lower availability of the case-marking cue in transitive sentences in German, as compared to Russian.

Subject-verb agreement. Subject-verb agreement is the other morphological cue that can provide reliable evidence for agentivity. In both Russian and German, the verb agrees in number and person with the subject of the sentence. In Russian, singular nouns and past tense verbs also agree in gender. The fact that Russian requires agreement on more morphosyntactic dimensions reflects a generally greater reliance on the agreement cue in Russian, as compared with German.

Corpus-based estimations. In order to quantify our assumptions regarding the differential availability of morphological cues in Russian and German, we extracted samples of active transitive sentences from a variety of sources and coded the presence or absence of each of these cues.

[1] End-palatalisation of consonants in Russian is graphemically expressed by the soft-sign which is transcribed by an apostrophe. In the transcription of Russian we follow the transliteration rules used by Comrie and Corbett (1992).

For each language, a corpus of 250 sentences was selected from contemporary novels, newspaper editorials, and children's books. The references are listed in Appendix 1. For each source, several pages were selected at random from which all sentences containing a single transitive verb were coded. In Russian, negative sentences were excluded because these sentences permit both accusative and genitive marking of the direct object. We also excluded sentences with null subjects. Since subject omission is more frequent in Russian, it was necessary to consult a somewhat larger text corpus in order to find a matching number of qualifying sentences.

We are aware of the fact that written language is not necessarily an accurate reflection of the distributional characteristics of the entire language. Unfortunately, large samples of adult oral speech are not yet available.[2] In an oral corpus, we would expect more word order variability. However, since the variation between SVO and OVS is determined by pragmatic or discourse factors in both languages, this cue should affect both languages to a similar degree. We would also expect more subject ellipsis, particularly in Russian. However, the omission of the subject does not affect the reliability or availability of the subject–verb agreement cue. There is little reason to believe that animacy effects are different in the two cultures. Finally, the reliability and availability of the case-marking cue is based on distributional effects that should be similar in written and oral corpora. Although written corpora give us a good initial idea of the relative availability and reliability of these cues, it is clear that, for future research, it would be desirable to base availability and reliability estimations on corpora of oral speech as well.

Since we are interested in the immediate effects of cues during sentence comprehension, animacy information and case-marking were coded as local cues for each individual noun. The ten specific cues that we tracked were:

1. animacy of the first noun (N1-Animacy)
2. inanimacy of the first noun (N1-Inanimacy)
3. animacy of the second noun (N2-Animacy)
4. inanimacy of the second noun (N2-Inanimacy)
5. nominative marking of the first noun (N1-Nominative)

[2] A good example of the use of oral corpora in psycholinguistic research can be found in the CHILDES database of child language and adult second language samples (MacWhinney, 1995), where spoken interactions from over 25 languages, including Russian and German, are available. Unfortunately, crosslinguistic studies of adult language production have not developed a similar focus on the collection, transcription, and analysis of spoken corpora. This is clearly a priority for future development in this field.

6. nominative marking of the second noun (N2-Nominative)
7. accusative marking of the first noun (N1-Accusative)
8. accusative marking of the second noun (N2-Accusative)
9. verb agreement with the first noun (N1-Agreement)
10. verb agreement with the second noun (N2-Agreement).

Additionally, the word order (NVN, NNV, VNN) of the sentence was coded. Availability of a cue was calculated as the proportion of codeable sentences in which the cue was present.

Reliability of the cue was calculated as the proportion of times the cue was associated with the first noun as agent, given that the cue was present (i.e. p Agent | Cue). Thus, a reliability of 1.0 indicates that, whenever the cue is present, it always points to the first noun as the agent. A reliability of 0.0 indicates that a cue always points to the second noun as the agent. A reliability of 0.5 indicates that the cue has no value at all in terms of helping us to decide whether the first noun is the agent. Both values close to 0.0 as well as values close to 1.0 are indicators of cues with extremely high reliability. The availability and reliability estimations for the various cues in both languages are given in Table 1.

As can be seen from Table 1, there are only minor differences between Russian and German in the reliability of the cues. When they are present and contrastive, the morphological cues of case-marking and subject–verb agreement are absolutely reliable. The reliability of animacy *vs* inanimacy

TABLE 1
Cue Availability and Cue Reliability in German and Russian

	Availability		*Reliability*	
	Russian	*German*	*Russian*	*German*
NVN	0.722	0.392	0.897	0.885
NNV	0.204	0.465	0.720	0.798
VNN	0.074	0.143	0.500	0.943
N1-Animacy	0.779	0.776	0.931	0.968
N1-Inanimacy	0.221	0.224	0.481	0.454
N2-Animacy	0.377	0.424	0.641	0.673
N2-Inanimacy	0.623	0.576	0.947	0.986
N1-Nominative	0.775	0.445	1.000	1.000
N2-Nominative	0.143	0.078	0.000	0.000
N1-Accusative	0.078	0.029	0.000	0.000
N2-Accusative	0.373	0.212	1.000	1.000
N1-Agreement	0.586	0.294	1.000	1.000
N2-Agreement	0.114	0.049	0.000	0.000

The reliability indicates how often a cue points to the first noun as the agent. NVN, noun-verb-noun, etc.

of the first and second noun is very similar in both languages and reflects the fact that animacy information is a semantic cue, rather than a language-specific cue. The reliability of word order is also quite similar in Russian and German. Despite the differences in permitted word orders, both languages have a strong tendency for the first noun to be the agent. Calculated over all word order types, the proportion of sentences where the first noun was the agent is 0.832 in Russian and 0.853 in German.

Russian and German differ primarily in terms of cue availability. While there were only small differences with respect to the availability of the animacy cue, the corpus analyses indicate that the availability of case-marking and subject–verb agreement is much higher in Russian than in German. Moreover, this pattern holds for both SO and OS configurations. The overall proportion of sentences without any morphological cue was 0.025 in Russian and 0.223 in German. The corpus-based availability and reliability estimations, therefore, suport the claim that Russian and German differ markedly with respect to the availability of morphological cues.

STUDY 1: SPEEDED PICTURE CHOICE EXPERIMENT

Our first study used the speeded picture choice task to examine possible language differences in on-line processing of the cues for agentivity. Since the reliability of these cues does not differ between Russian and German, we do not expect significant differences in the distribution of agent choices. However, we expect to find differences in reaction times due to differential effects of cue availability.

A complete crossing of all available cues described in the corpus analyses would result in 288 grammatically correct sentence types for each language. Since testing all these conditions is not feasible within one study, we confined the design to a systematic crossing of animacy, subject–object configuration, and case-marking in NVN sentences. If the lower availability of case-marking in German has an effect on on-line processing, we would expect the reaction time benefit for sentences with case-marked nouns as compared to sentences with unmarked nouns to be larger in Russian than in German. Moreover, the Competition Model leads us to expect that any such weakness in the use of the case-marking cue in German should be matched by a compensatory reliance on animacy information. If this is the case, then reaction time benefits from converging animacy information should be larger in German than in Russian.

Method

Participants. Participants were 14 native speakers of German and 14 native speakers of Russian. The speakers of German were studying or working in the United States. Seven of the Russian participants were immigrants and the remaining seven were enrolled in a student exchange programme or were relatives of exchange students. All but one participant in the Russian group had received partial or complete college or university education in their native language. None of the participants had lived in an English-speaking environment for longer than 2 years. The participants' exposure to other languages was assessed through a background questionnaire which indicated that none of the participants had been exposed to any of their second languages before the age of 11. The mean age of the participants was nearly identical in both groups (Russian: 25.7 years; German: 26.6 years). All participants were paid $5.00 for their participation.

Materials and Design. Language was varied as a between-subjects factor. The stimulus materials consisted of simple active transitive NVN sentences which were grammatically correct in both Russian and German. Within each language, the sentences were varied according to the following five factors:

1. animacy of the first noun (N1-Animacy: animate *vs* inanimate)
2. animacy of the second noun (N2-Animacy: animate *vs* inanimate)
3. configuration (SVO *vs* OVS)
4. case-marking of the first noun (N1-Marking: marked *vs* unmarked)
5. case-marking of the second noun (N2-Marking: marked *vs* unmarked)

Appendix 2 presents the composition of the 32 cells of this design. The configuration factor specifies the type of marking on the two nouns: for SVO sentences, case-marking on the first noun was nominative and on the second noun accusative. In OVS sentences, the pattern was the opposite: case-marking on the first noun was accusative and on the second noun nominative. Language was varied as a between-subjects factor. When case-marking is neutralised (unmarked-unmarked), it is impossible to distinguish between the SVO and OVS configurations and listeners will normally impose an SVO interpretation. This neutralisation occurs for four cell pairs in Appendix 2: 1 and 5, 9 and 13, 17 and 21, and 25 and 29. When interpreting the results, it is important to remember that sentences with unmarked case in the OVS condition will be interpreted as SVO.

The sentences were composed from combinations of four animate nouns, four inanimate nouns, and two verbs. In order to minimise semantic differences between languages, exact translation equivalents for all ten words were used. The English translations of the eight nouns were 'mother', 'daughter', 'father', 'son', 'flower', 'plate', 'spoon', and 'cake'. The verbs were 'looking for' and 'finding'. In German, the nouns 'mother', 'daughter', 'flower', and 'cake' are feminine and exhibit nominative-accusative neutralisation. In Russian, there is also nominative-accusative neutralisation for these four nouns because 'mother' and 'daughter' have final palatalisation and because 'flower' and 'cake' are masculine inanimate. Since all nouns were singular and all verbs in present tense, verb agreement was not available as a cue.

For each condition, four sentences were constructed by first combining the relevant nouns in all possible ways and then counterbalancing the two verbs. This resulted in 128 grammatically correct sentences. Although all of the sentences were grammatical, half of them were semantically implausible. Implausible sentences correspond to English sentences such as 'The cake finds the father'. In these sentences, animacy is in competition with case-marking and/or configuration. Examples for all 32 sentence types are also presented in Appendix 2.

The eight nouns and two verbs were recorded by a female native speaker and digitised with a 16-bit sampling rate at 22 kHz using SoundEdit16. Combining single word recordings into sentences ensured that the intonation pattern was identical for all sentences and that no prosodic cues were available to the listener. The presentation of the digitised auditory stimuli and of the eight pictures depicting the nouns in the sentence were controlled by the PsyScope experimental control programme (Cohen, MacWhinney, Flatt, & Provost, 1993). The pictures of the four inanimate nouns were taken from the Snodgrass and Vanderwart (1980) materials. The pictures of the four animate nouns were taken from other sources.

Procedure. The participants were familiarised with the pictures and the nouns by presenting each picture one at a time in the middle of the screen, along with the correct label for each picture. Next, participants saw pairs of pictures. As soon as the pictures were displayed, the label of one of the two pictures was presented through headphones. Participants placed their left and right index fingers on the two outer buttons of the CMU button box and were given these instructions: "As fast as possible, press the button closer to the object being named." Thus, if the left picture was named, the left button had to be pressed and vice versa. The purpose of this preparation phase was to acquaint the participants with the speeded forced choice task and to familiarise them with the specific picture-noun combinations. Participants saw all 54 possible combinations of the eight pictures in individually randomised order.

In the main experiment, participants were told that they were going to hear a series of simple transitive sentences accompanied by the pictures of both nouns. Again, their task was: "As fast as possible, press the button closer to the picture that is depicting who or what was carrying out the action." If the left picture depicted the agent, the left button had to be pressed and vice versa. Participants were also told that some of the sentences might be implausible and that they should make a choice regardless of whether the sentence made sense to them or not. It was stressed that they had to make a choice as soon as they felt certain enough, even if they had not yet finished hearing the sentence.

For each participant, ten practice sentences were selected randomly from the larger pool of sentences. After the ten practice trials, all 128 sentences were presented. Order of presentation was randomised individually by the PsyScope program. A trial consisted of the presentation of a fixation point in the middle of the screen for 1000 msec followed by simultaneous presentation of the sentence and the two pictures. After the participants had executed their response, the next trial followed with an ISI of 500 msec. Participants' choice and reaction times were recorded. The experiment was carried out on a Macintosh Centris 660 AV.

Results

Choice Responses. Table 2 presents the choice data for all conditions as proportions of first noun choices. Although the first noun is always the correct choice in case-marked SVO sentences and the second noun is always the correct choice in case-marked OVS sentences, we do not present the choice data as errors. This is because, in the morphologically ambiguous sentences, there was no reliable cue for determining a correct interpretation.

The arcsin-transformed proportions of first noun choices for each language group were submitted to an omnibus ANOVA. The first column of Appendix 3 lists all effects that reached or almost reached significance at the $P < 0.05$ level both in the analyses by subjects (F_1) and by items (F_2). This analysis revealed no significant language differences. The large main effect of Configuration indicates that, both in Russian and in German, agent choice was determined by case-marking. When case-marking was available, participants chose the first noun in SVO sentences and the second noun in OVS sentences. This main effect is specified by the 2-way interactions of Configuration with N1-Marking, Configuration with N2-Marking, and N1-Marking with N2-Marking, as well as the 3-way interaction of Configuration, N1-Marking, and N2-Marking which is depicted by the solid lines in Fig. 1.

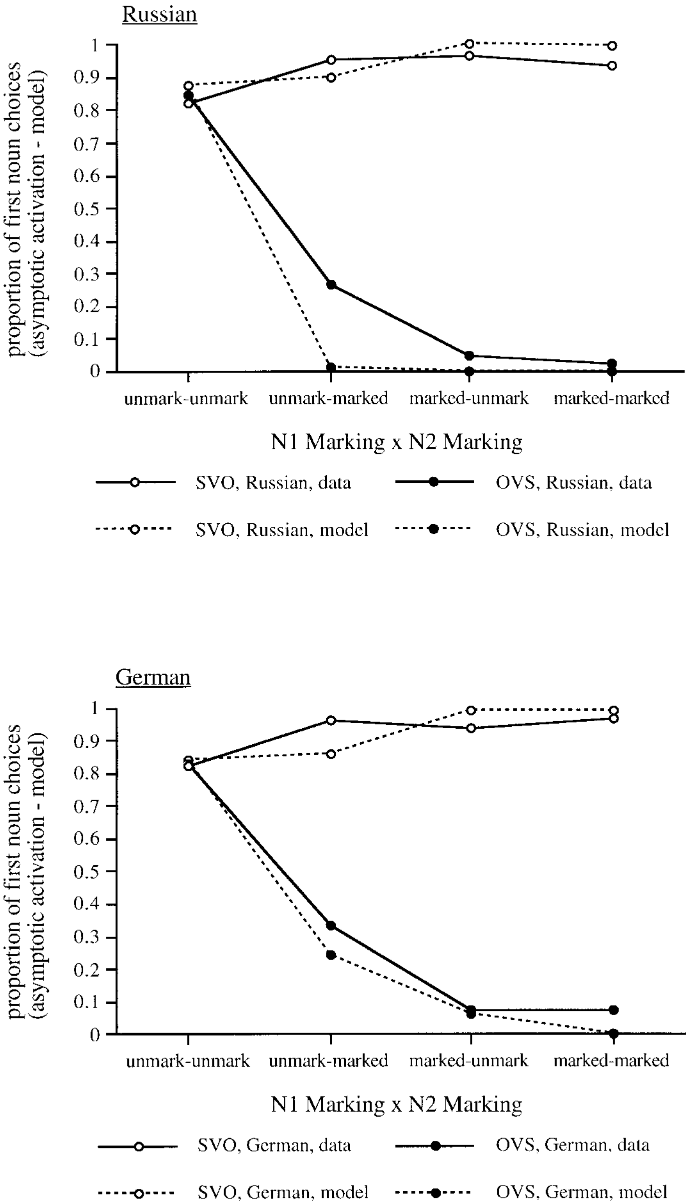

FIG. 1. Proportions of first noun choices in the experimental data (solid lines) and asymptotic activation in the model (dotted lines) as a function of Configuration, N1-Marking, and N2-Marking for Russian (upper panel) and German (lower panel).

TABLE 2
Proportion of First Noun Choices per Condition in the Human Data and in the Model

					Russian		*German*	
N1-Animacy	*N2-Animacy*	*Configuration*	*N1-Marking*	*N2-Marking*	*Data*	*Network*	*Data*	*Network*
anin	anim	SVO	unmark	unmark	0.86	0.95	0.93	0.97
				marked	1.00	1.00	0.98	0.99
			marked	unmark	0.98	0.99	1.00	0.98
				marked	0.98	1.00	0.97	0.99
		OVS	unmark	unmark	0.92	0.95	0.93	0.97
				marked	0.27	0.02	0.32	0.08
			marked	unmark	0.02	0.00	0.04	0.00
				marked	0.02	0.00	0.07	0.00
	inanim	SVO	unmark	unmark	0.94	0.99	0.97	0.98
				marked	0.97	1.00	0.98	0.99
			marked	unmark	1.00	0.99	1.00	1.00
				marked	0.95	1.00	0.96	1.00
		OVS	unmark	unmark	0.92	0.99	1.00	0.98
				marked	0.28	0.02	0.45	0.80
			marked	unmark	0.06	0.00	0.10	0.03
				marked	0.03	0.00	0.09	0.00
inanim	anim	SVO	unmark	unmark	0.58	0.78	0.55	0.50
				marked	0.88	1.00	0.85	0.99
			marked	unmark	0.86	0.89	0.86	0.48
				marked	0.95	1.00	0.94	0.99
		OVS	unmark	unmark	0.66	0.78	0.56	0.50
				marked	0.25	0.01	0.18	0.00
			marked	unmark	0.01	0.00	0.06	0.00
				marked	0.02	0.00	0.02	0.00

TABLE 2 *(Contd)*
Proportion of First Noun Choices per Condition in the Human Data and in the Model

					Russian		*German*	
N1-Animacy	*N2-Animacy*	*Configuration*	*N1-Marking*	*N2-Marking*	*Data*	*Network*	*Data*	*Network*
	inanim	SVO	unmark	unmark	0.89	0.77	0.83	0.92
				marked	1.00	1.00	0.96	0.99
			marked	unmark	0.97	0.74	0.98	0.97
				marked	0.97	1.00	1.00	1.00
		OVS	unmark	unmark	0.88	0.77	0.81	0.92
				marked	0.27	0.00	0.39	0.10
			marked	unmark	0.09	0.00	0.07	0.20
				marked	0.03	0.00	0.12	0.00

The fourth and fifth columns in Appendix 3 display the percent of experimental variance accounted for by each effect or interaction in both the analysis by subjects and by items. Together, the various effects involving case-marking and configuration accounted for approximately 93% of the experimental variance. The participants chose the first noun 82% of the time when sentences were not marked for case, as compared to almost 100% of the time in case-marked SVO sentences. The fact that, in the absence of case-marking, the first noun is chosen above chance suggests that, in both languages, there is a default bias towards interpreting the first noun as the agent if no additional cues are available. This bias is, however, rapidly reversed towards a correct second noun choice if an accusative marker occurs on the first noun. In sentences with unmarked first nouns nd nominative-marked second nouns, the initial 'first noun' interpretation has time to become fairly well consolidated by the time the listener encounters the conflicting nominative case-marker later in the sentence. This leads to a substantial amount of erroneous first noun choices. The average proportion of erroneous first noun choices in this condition was 0.27 in Russian and 0.34 in German. In both languages, the proportion of errors in nominative-marked OVS sentences is significantly higher than in accusative-marked OVS sentences (all ts > 9.4, all Ps < 0.001).

The main effect of N1-Animacy indicates that animate first nouns are more readily interpreted as agents. This tendency is, however, easily overridden by case-marking, as the interaction of N1-Animacy and N1-Marking indicates (see Fig. 2, solid lines). The effects of N2-Animacy was relatively weak indicating that the second noun had much less influence on the decisions than the first one.

Reaction Times. All outliers above 4000 msec were truncated, thereby excluding 3.5% of the data points. Since all sentences were grammatically correct, we included only correct responses into the statistical analysis. For case-marked sentences, correct responses were those indicated by the nominative and/or accusative case-markers. For unmarked sentences, first noun responses were defined as the correct ones, reflecting the fact that SVO is the canonical, unmarked order in both languages. This is justified by the finding that both language groups exhibited a strong first noun bias in these sentences. Overall, another 12% of the data points were excluded by these criteria. Due to the exclusion of these data points, 13 condition means per subject were missing and had to be substituted following a procedure recommended by Winer (1971, p. 487).

The analysis of the reaction times is complicated by the fact that the auditory presentation of the stimulus material induced an additional source of variability due to the different duration of the various parts of the sentence. For example, Russian animate nouns were always 154 msec

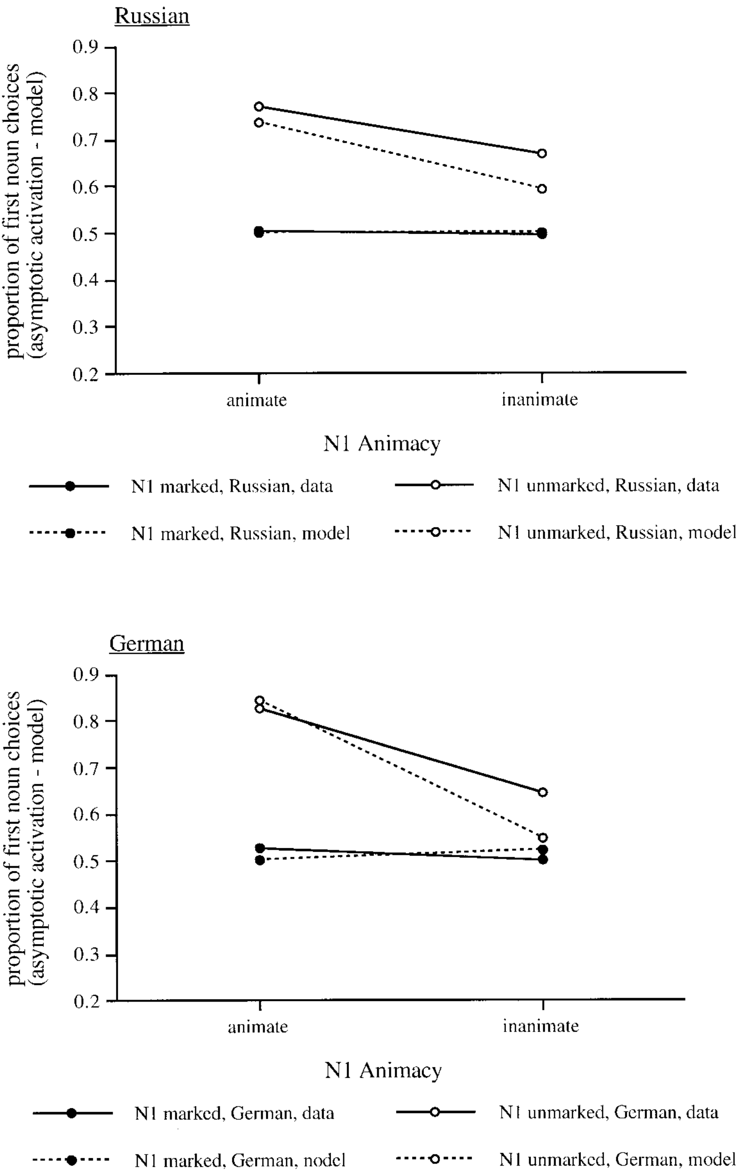

FIG. 2. Proportions of first noun choices in the experimental data (solid lines) and asymptotic activation in the model (dotted lines) as a function of N1-Marking and N1-Animacy for Russian (upper panel) and German (lower panel).

shorter than Russian inanimate nouns, as opposed to a difference of only 77 msec in German. Similarly, some of the Russian accusative marked masculine animate nouns were longer than their unmarked nominative form because the addition of the suffix changes the syllabic structure of these nouns [e.g. *syn* (NOM)—*syna* (ACC)]. Moreover, the languages differed with respect to the position of the case markers within the sentences. In German, case-marking affects the article which precedes the noun whereas in Russian, it affects the suffix at the end of the noun. This also has consequences for how fast a decision about the agent of the sentence can be reached. Figure 3 illustrates the differences in the temporal structure of a typical Russian and German experimental sentence by displaying the position of the case-markers in relation to the average duration of the two nouns and the verb. These duration differences need to be accounted for, if reliable conclusions about the on-line effects of the various cues are to be drawn, particularly in regard to the language differences.

In order to account for the duration differences, we adjusted the reaction times to various reference points in the sentences and performed the statistical analyses on these adjusted reaction times. First, we adjusted the reaction times to the end of the first noun by subtracting the duration of the first noun from the reaction time for each sentence. This adjustment controls for potential effects due to differences in the duration of the first nouns. Second, we adjusted the reaction times to the end of the verb thereby controlling for duration differences of the first noun and the verb combined. Finally, we adjusted the reaction times to the end of the sentence which additionally accounts for duration effects due to the second

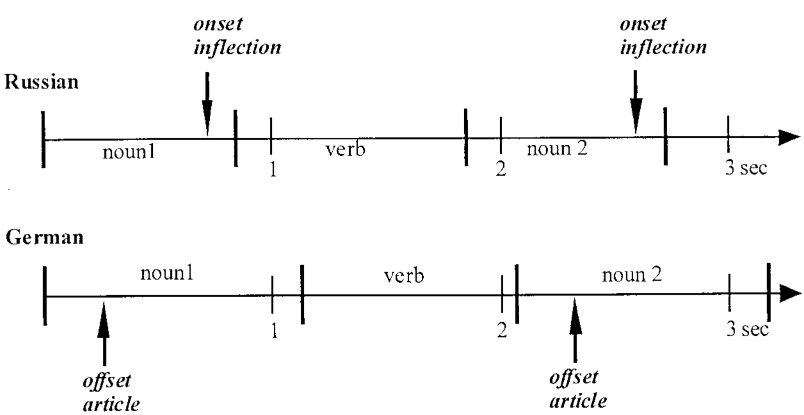

FIG. 3. Average duration of the noun phrases and the verb as well as positioning of case-markers in the Russian and German sentences.

noun. Omnibus ANOVAs by subjects and by items were performed on all three adjusted reaction times. Unless specified in the text, all effects interpreted below are statistically reliable for all three reaction time adjustments. This restriction is necessary since effects that are not similar across adjusted reaction times may be confounded with differences in the duration of the various parts of the sentence.[3] The corresponding *F*-values and percentages of experimental variance accounted for by the various effects and interactions are listed in Appendix 4.

We will first describe the effects that were similar in both language groups, i.e. effects that did not interact with the factor of Language. The strongest effect in both language groups was the effect of case-marking on the first noun which accounted for about 40% of the experimental variance. If the first noun was case-marked, reaction times were on average 318 msec faster than if it was unmarked. This effect was specified by an interaction of N1-Marking and Configuration which is due to the fact that the average reaction time benefit of 499 msec from accusative marking on the first noun in OVS sentences is larger than the average benefit of 151 msec from nominative marking on the first noun in SVO sentences.

There was also a significant, albeit weaker, effect of case-marking on the second noun. This effect was further specified by a series of 2-way and 3-way interactions with N1-Marking and Configuration. These interactions are due to longer reaction times in sentences with unmarked first and nominative marked second nouns of the type *Die Tochter sucht der Vater.* in German and *Doč iščet otec.* in Russian as compared to unmarked sentences like *Die Tochter sucht die Mutter.* or *Doč iščet mat'.* For these sentences, an additional 327 msec were required to restructure the initial interpretation of the first noun as the agent, as soon as the unexpected nominative marker on the second noun is encountered. The initial incorrect interpretation was further strengthened if the first noun was animate, as indicated by the significant 3-way interaction of Configuration, N1-Animacy, and N2-Marking. However, case-marking of the second noun had no additional effect if the first noun was unambiguously case-marked.

Next, we will describe those effects that were different in both language groups as indicated by interactions with the factor of Language. First, there

[3] The use of ANCOVAs to partial out the effects of word duration is not a viable alternative because the assumption of constant slopes is violated. Thus, the duration of the various sentence parts has different effects in the two languages. For example, the duration of the nouns has much more of an effect in Russian because the case marker appears at the end of the noun so that the listener will be much more inclined to process the whole noun, as opposed to German, where decisions can be made right after the article and before the whole noun has been processed.

was a significant interaction between Language and N1-Marking in the analyses by items. However, this effect fell short of significance in the analyses by subject ($P = 0.06$ for reaction times adjusted to the end of the first noun and $P = 0.08$ for reaction times adjusted to the end of the verb and the second noun). This interaction indicates that the reaction time benefit of 400 msec from first noun marking in Russian was larger than the benefit of 228 msec in German. Note that this interaction was found in reaction times adjusted to the end of the nouns which did not account for relative duration differences between articles and nouns in the various conditions in German. Still, the relative benefit from case-marking on the first noun was larger in Russian despite the fact that the average duration of the German case-marked nouns was slightly shorter in terms of number of syllables. This suggests that the language differences in the processing benefit from case-marking cannot be attributed to word duration effects. Furthermore, the interaction between Language and N1-Marking was specified by a 3-way interaction involving the factor Configuration which indicates that the language differences in the benefit from first noun-marking were most pronounced for accusative marked first nouns. The upper panels of Fig. 4 depict this interaction in the reaction times adjusted to the end of the first noun.

The other important reaction time difference between Russian and German was related to the effects of animacy of the first noun. This is suggested by the interaction between Language and N1-Animacy, which reached significance in four out of the six ANOVAs and fell short of significance in the analyses by items for reaction times adjusted to the end of the verb ($P = 0.08$) and to the end of the second noun ($P = 0.09$). In the German group, but not in the Russian group, decisions are faster when the first noun is animate. Apparently, the relatively weaker validity of nominative-marking in German leaves more room for animacy information to impact on-line performance. We also found a 3-way interaction between Language, Configuration, and N1-Animacy in the analyses by items. However, this effect fell short of significance in the analyses by subjects ($P = 0.1$ for reaction times adjusted to the end of the first noun; $P = 0.07$ for reaction times adjusted to the end of the verb and the second noun). This suggests that the N1-Animacy benefit observed in German seems to be confined to SVO sentences, in which the presence of an animate first noun serves to further support the ‘agent first’ bias. Apparently, in OVS sentences, accusative-marking on the first noun tends to override any effects of N1-Animacy in both languages, thereby supporting the earlier finding that accusative markers are stronger cues than nominative markers. The upper panels of Fig. 5 depict this interaction in the reaction times adjusted to the end of the first noun.

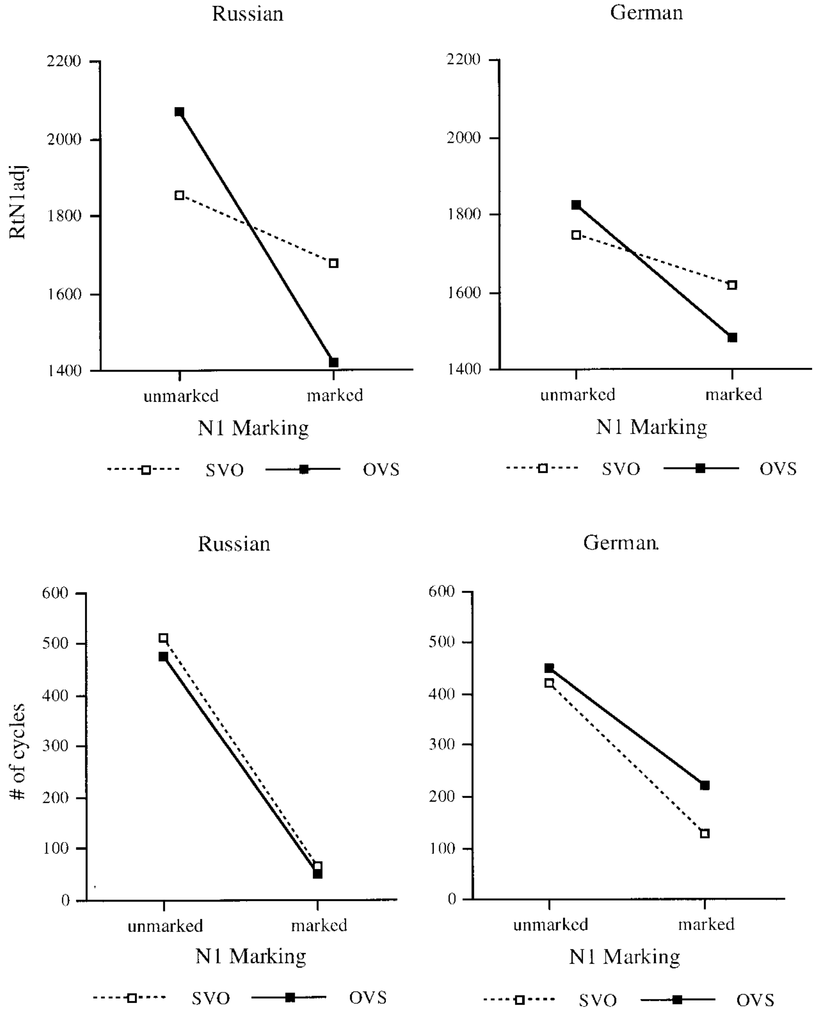

FIG. 4. Reaction times adjusted to the end of the first noun in experiment 1 (upper panels) and in the reaction time estimates derived from the model (lower panels) as a function of nominative-marking (SVO) and accusative-marking (OVS) in Russian and German.

Discussion

The off-line results of the speeded agent choice task demonstrate that, in many regards, speakers of Russian and German behave similarly in their interpretation of the sentences. In the vast majority of trials, the final interpretation is determined by the most reliable cue, which is case-

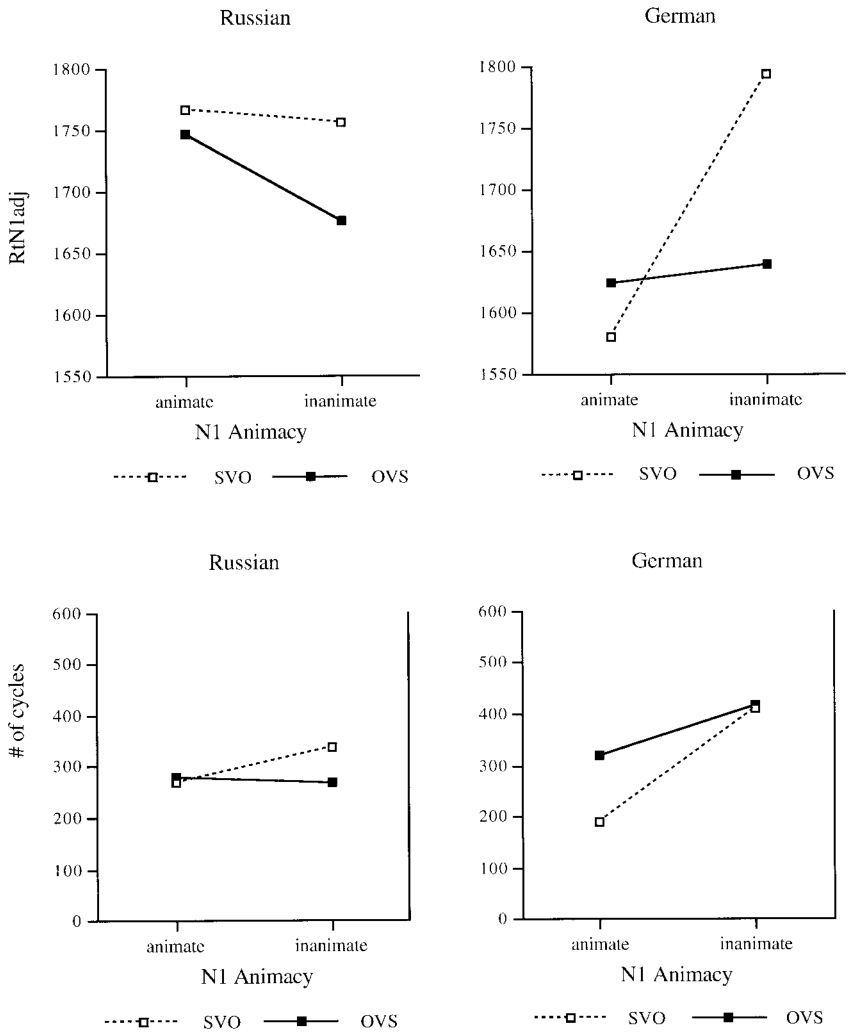

FIG. 5. Reaction times adjusted to the end of the first noun in experiment 1 (upper panels) and reaction time estimates derived from the model (lower panels) as a function of N1-Animacy and Configuration in Russian and German.

marking. This comes as no surprise. In both languages, we found that the animacy cue had a greater effect on speeding up reaction times when it occurred on the first noun, than when it occurred on the second noun. This effect matches up well with the idea that the first noun plays a unique role in organising and grounding the conceptual structure of the rest of the sentence (MacWhinney, 1977; Gernsbacher Hargreaves, & Beeman, 1989).

However, these findings regarding similarities between the two languages were supplemented by some interesting language differences in on-line processing which had been predicted by our extended version of the Competition Model. Specifically, the corpus analysis predicted that case-marking should be a weaker cue in German than in Russian. This means that, in relative terms, animacy should be a stronger cue in German than in Russian. Two findings confirmed these predictions. First, the magnitude of the processing benefits from case-marking on the first noun was larger in Russian, supporting the idea that case-marking in general is a stronger cue in Russian than in German. Both the processing benefit of first noun case-marking and the language difference in the magnitude of this effect were most pronounced for accusative-marking. This aspect of the result is somewhat puzzling. The smaller effect of nominative-marking as compared to accusative-marking did not fall out obviously from the corpus-based availability estimations. Collapsed over all nouns, the difference in availability of the nominative and the accusative was of similar size for Russian and German. Looking only at reliability and availability, we cannot account for the stronger effect of accusative case-marking. Second, the reaction time effects attributable to animacy of the first noun were different in Russian and German. In Russian, the strong use of nominative-marking swamps any effects of animacy of the first noun. Basically, speakers of Russian look for first noun case. If there is unambiguous case-marking, it is used rapidly, ignoring any additional information. In German, however, first noun animacy leads to similar reaction time benefits for both unmarked and nominative-marked first nouns, thus indicating that speakers of German take both case and animacy cues into account when processing the first noun.

The fact that all sentences in the experiment were grammatical and did not contain unusual cue combinations supports the ecological validity of the task and the generalisability of this interpretation. It can be argued, however, that the speeded nature of the task might have induced unusual performance patterns. In this regard it is noteworthy that the analyses of the choice data revealed no language differences while the analyses of the reaction times did. This suggests that the obtained processing differences between the two languages cannot be attributed to task-induced strategies or to a trade-off between speed and accuracy, but reflect basic differences in automatised aspects of sentence processing.

Finally, the results revealed an interesting pattern of cue interaction. The interaction between case-marking of the first and second noun revealed that a second case-marker leads to no additional benefit if the first noun is already reliably case-marked. This finding replicates results obtained by Kail (1989) and Mimica et al. (1994) with respect to the convergence of cues, which have shown that, in the presence of a strong cue, a weaker supporting cue might show no additional effect. In the situation of a speeded task, this can lead to early decisions that are made right after the encounter of the strong cue and before all other cues in a sentence have been fully processed. These non-additive patterns of cue interaction cannot be predicted from cue validity estimations for each of the cues separately.

Taraban and Roark (1996) have shown that the performance of a backpropagation learning model showed a better fit to the patterns of cue interaction than estimations based on the reliability of individual cues. They have argued that connectionist models provide a more accurate way to account for effects of cue competition and cue convergence. The next section describes an attempt to extend this approach towards cue interaction in on-line processing. We will try to simulate the obtained empirical results using a recurrent cascaded backpropagation network. The aim of this simulation is to explore to what degree the language differences and patterns of cue interaction in the reaction times can be accounted for by principles of parallel interactive activation. In evaluating the network performance, we will concentrate on the following five main results from the experiment:

1. *Availability effects.* The discrepancy between choice data and reaction times showed that differences in cue availability tend to affect the time course, but not the final result of processing.
2. *Case-marking effects.* The case-marking cue has an immediate effect which is stronger in Russian than in German. This language difference has been attributed to the higher availability of this cue in Russian.
3. *Animacy effects.* The language differences in the availability of case-marking result in reliance on the animacy cue in German, but not in Russian.
4. *On-line cue interaction.* The stronger an early cue, the more it tends to override later cues in on-line processing without displaying any patterns of information integration. On the other hand, the weaker a cue, the more does processing rely on other supporting cues.
5. *Accusative superiority.* There was an overall stronger benefit from accusative-marking as compared to nominative-marking in the reaction times. This effect was not predicted by the corpus-based validity estimations.

STUDY 2: CONNECTIONIST SIMULATION

The Competition Model notions of availability and reliability have been useful as guides to off-line choice behaviour. They also work to predict certain basic aspects of language differences in on-line processing. However, they do not offer a satisfactory account for the specific patterns of cue interaction. The connectionist model developed in this section attempts to go beyond these two rather simplistic constructs and to formulate a fuller account of on-line processing effects. The model presented here is designed to capture basic language differences in the cue distributions. It is not presented as an attempt to model the full structure of either Russian or German. Instead, it uses a simple architecture and a limited training corpus to model patterns of cue interaction in the reaction time data of Study 1.

The model emphasises the sequential temporal nature of the sentence interpretation task. This temporal sequentiality is captured by using a simple recurrent network (Elman, 1990) which takes in input one word at a time. The network keeps track of previous activation states, while new information comes in. As the input is presented, the activation of various alternative interpretations builds up gradually. Cascaded networks (McClelland, 1979) in which the activation of units is calculated as the running average of their net input over time have been shown to exhibit a good fit to reaction time data (Cohen, Dunbar, & McClelland, 1990). Combining the advantages of cascading with those of the recurrent network, we created a recurrent cascaded network to simulate the on-line effects of availability, reliability, and cue interaction. Learning occurs via the backpropagation rule (Rumelhart, Hinton, & Williams, 1986) which modifies connection strength as a function of feedback. The output of the model is compared to a target signal and the learning algorithm adjusts the connection weights in a way that reduces overall error. Back-propagation networks have been shown to be sensitive to the frequency of input patterns since frequent patterns make large contributions to the overall error reduction.

The network consisted of four layers: an input layer with five units, a context layer with five units, a hidden layer with five units, and a single output unit. The five input units coded the following properties of an input string: (i) phrasal word class (noun phrase *vs* verb phrase), (ii) animacy, (iii) nominative-marking, (iv) accusative-marking, and (v) a single coding for any of the morphosyntactic dimensions (gender or number) on which nouns and verbs can agree. If a string was coded as a verb, the animacy and case-marking units were always set to 0. The single output unit coded whether the first noun was the agent of the sentence. The values of the target output were set to 1 for a ‘first noun’-response and to 0 for a ‘second noun'-response. Thus, the activation strength in the output unit corresponds to the strength of the 'first-noun'-interpretation which can be directly compared to the proportion of first noun choices in the empirical data.

At each trial, the input corresponding to a sentence is processed in the following way: after the first string is presented, activation is fed forward and compared to the target. Using the backpropagation algorithm, all weights are changed in a way that minimises the overall error term. The activation state of the hidden units is then copied back to the context units and presented as input together with the second string. After processing the second string, the newly obtained activation state is copied back and presented as context to the input of the third string. When processing of three input strings is completed, the values of the context units are cleared. The architecture of the network is presented in Fig. 6.

The recurrent architecture of the network allowed us to match the distribution of cues over the course of the sentence. The local cues of animacy and case-marking were coded as features within an input string, whereas the topological cue of verb agreement was coded as a match of activation in the input units over time. For example, the input coding for the three strings in an SVO sentence containing two animate case-marked nouns and verb agreement was 11101/00001/11010. Agreement between the first noun and the verb is coded by virtue of the match of the value "1"

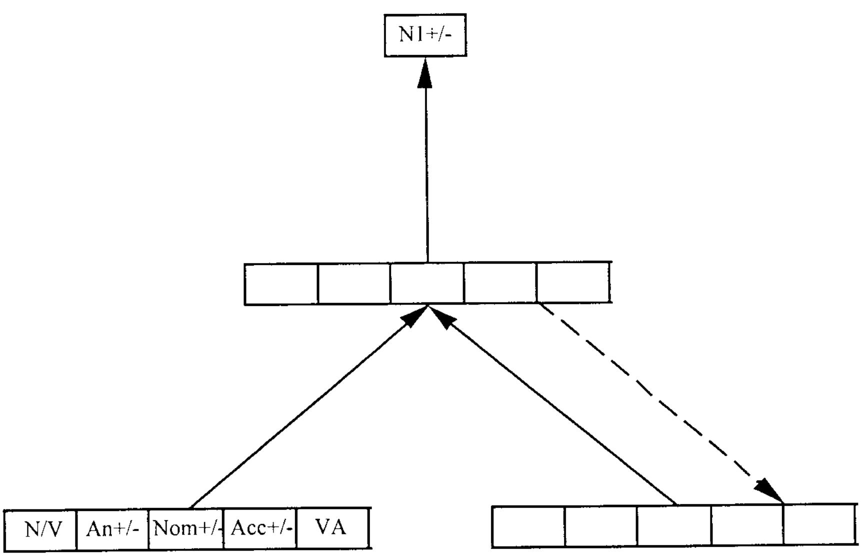

FIG. 6. Architecture of the recurrent cascaded backpropagation model.

in the final unit of the first and the second strings. In order to prevent the network from developing verb-agreement biases associated with a particular value of the unit, half of the patterns had agreement coded as a match in the value of 1 (as in the example above) and the other half as a match in the value of 0.

We constructed two sets of training patterns in order to model the essential differences in the cue distributions between Russian and German. Each pattern in the training sets corresponded to a specific sentence type. The sentence types varied on the dimensions of word order, animacy of the two nouns, case-marking of the two nouns, and verb agreement. The training sets resembled the frequencies of all sentence types obtained from the corpus analyses, thereby reproducing the essential language differences in the availability of the cues. The German corpus included 71 different sentence types; the Russian corpus included only 51 different sentence types. For the ambiguous sentence types, the agent was determined on the basis of the contextual information given in the corpus.

We used a very low learning rate of 0.01 and a momentum of 0.7. Furthermore, the proportion of activation of the hidden units that was copied back to the context units (μ) also had to be set to the fairly low value of 0.4 in order to allow for drastic activation changes induced by competing evidence presented at different points in the input patterns. Each network was trained for 3000 epochs on the Russian *vs* the German training patterns until the overall error term had reached an asymptotic value in both sets. During learning, the connection weights were updated after each sweep through the whole set of training patterns. For both languages, we ran five training simulations each with a different set of random weights at the outset of learning. After training, the learning performance was tested for each of the five runs in each language. The test patterns coded the 32 sentence types used in Experiment 1. Ten of the 32 test patterns had been present in the German training set, and 18 of the 32 test patterns had been present in the Russian training set. In the test phase, the network was presented with the 32 test patterns and was allowed to cycle until the output unit had reached an asymptotic value. The number of cycles for each input string was set to 200. It has been shown that the asymptotic values reached in cascaded feedforward networks are identical to the activation reached in a single step in standard one-pass computations (McClelland, 1979). However, the cascaded nature of the network allows us to examine the gradual build-up of activation.

In order to compare the network performance with the empirical data we saved the asymptotic activation values reached after presentation of each of the three input strings in a pattern. The asymptotic activation arrived at after the third string, which corresponds to the last segment of the sentence, was compared with the overall proportion of first noun choices for the corresponding sentence type obtained in Experiment 1. Furthermore, for each pattern, we determined the number of cycles necessary to reach a threshold which was taken as an estimation of the reaction time. The threshold values were set to 0.98 for a first noun response and to 0.02 for a second noun response. This threshold prevented us from coding an incorrect response in patterns where animacy contrast conflicted with case-marking on the second noun. Note that this threshold can be reached after the first string (within 0–200 cycles), after the second string (within 200–400 cycles), or after the third string (within 400–600 cycles), thereby accounting for the fact that, in a speeded choice task, responses can be given at any time during sentence comprehension. If the asymptotic activation did not reach the threshold at all, we coded the maximum number of 600 cycles as reaction time estimation. The next section evaluates the fit of the model to the human data with respect to the proportions of the first noun choices and the reaction times.

Fit to Choice Data. The mean asymptotic activation values calculated over the five runs in each language can also be found in Table 2. The overall fit of the model to the experimental data was determined through a regression of the mean activations onto the proportions of first noun choices. This analysis revealed an R^2 of 0.95 ($P < 0.001$) for Russian and 0.92 ($P < 0.001$) for German, thus exhibiting an excellent fit of the model to the human data. It demonstrates that even for this limited training set a recurrent backpropagation network is able to learn the correct interpretations of a wide variety of Russian and German sentence patterns, even if they were not present in the learning input.

In order to better evaluate the extent to which the network reproduced the qualitative profile of the choice data, we submitted the arcsin-transformed asymptotic activations for the five runs per language to an omnibus ANOVA. Since the variance over the five runs was quite small many of the factorial effects reached significance. Therefore, we will not present the F-values but rather compare the effect sizes of the model and the human data in terms of percentage of variance accounted for by each effect and interaction. The model and human effect sizes, calculated as the percentage of experimental variance accounted for by each effect, are presented in Appendix 3.

As in the human data, no significant language differences were found. The hierarchy of the effect sizes is the same in the network and in the human data: the largest effect size was found for the factor of Configuration (about 61%) followed by the interaction of N1-Marking and Configuration (about 14%) and of N2-Marking and Configuration (about 5%). This indicates that in the network, as in the human data, noun choice is overwhelmingly determined by case-marking. Furthermore, the analysis revealed interactions between first noun marking and first noun animacy confirming that, as in the human data, case-marking overrides animacy information in the network's final result of sentence interpretation. The dotted lines in Fig. 1 show the model performance with respect to the effects of Configuration, N1-Marking, and N2-Marking and in relation to the human data. As can be seen, the direction of the effects as well as the overall effect sizes of model and data are almost identical.

One source of minor discrepancies between the learning outcome of the network and the human choice data are patterns with unmarked first nouns and nominative-marked second nouns (see Table 2). Recall that these patterns correspond to the sentences requiring a restructuring of the initial 'agent first'-interpretation. Here, the Russian network exhibited more accurate performance than the speakers of Russian. This discrepancy is most likely a result of the speeded nature of the task which caused the participants to terminate processing before they had settled into a stable response. In contrast, the network always reaches asymptotic activation. While the performance of the German network was also superior for most of these patterns, the speakers of German outperformed the network when second noun nominative-marking conflicted with the animacy contrast in AI sentences. For this pattern, the erroneous first noun activation of the network (0.80) was clearly higher than in the human data (0.45) (see Table 2), which indicates that the German network weighted the animacy cue somewhat higher than the Russian network. This suggests that the effect of noun animacy was a bit exaggerated in the German network. However, the fact that this cue is weighted higher in German is in accordance with the stronger reliance on animacy found in the German reaction times. Taken together, the learning outcome of the network revealed a remarkable fit to the choice data of the experiment.

Fit to Reaction Times. In order to evaluate the general fit of the reaction time estimations derived from the model, we determined how much variance in the human reaction times is accounted for by the number of cycles necessary to reach the threshold. Because the reaction times in the experimental data are influenced by the actual duration of the words in milliseconds, we need to partial out these stimulus effects in order to properly evaluate the match of the model to the data. To do this, we performed a stepwise regression on the mean reaction times per sentence type with duration of the first and the second noun entered at the first step and number of cycles entered at the second step. Duration of the verb was not entered since the counterbalancing scheme used for constructing the experimental materials had resulted in identical mean verb duration per sentence type. The percentage of variance accounted for by the number of cycles above the duration of the first and the second noun was 24% for

Russian [$F(1,28) = 11.6$, $P < 0.01$] and 38% for German [$F(1,28) = 39.3$, $P < 0.001$]. Thus, the network performance accounted for a substantial percentage of variance in the human data. The general fit of the model was somewhat better for German.

An omnibus ANOVA was performed on the number of cycles in the five runs in order to evaluate the qualitative profile of the network performance with respect to the time course of processing. The effect sizes are presented in the last column of Appendix 4. We will compare the performance of the network to the human data with respect to the language differences in on-line processing of case-marking and noun animacy. First, as in the human data, the largest proportion of variance was accounted for by the effect of N1-Marking. The average difference in the number of cycles between unmarked and marked first nouns was 463 in the Russian and 314 in the German network. However, the overall magnitude of this effect is exaggerated in the network where it accounted for 81.4% of variance as compared with only an average of 42% of variance in the human data. This effect was further specified by the interaction of Language and N1-Marking which indicates that the magnitude of the N1-Marking benefit was larger in the Russian network than in the German network. Thus, the network captures the essential language difference in the effect of case-marking on the first noun. However, a comparison between the upper and lower panels of Fig. 4 demonstrates that the network did not differentiate between effects of first noun-nominative-marking and first noun accusative-marking, as evidenced in the absence of any interaction of N1-Marking with Configuration in the network performance. This suggests that the larger benefit from accusative-marking cannot be explained by the distributional characteristics of cues in Russian and German transitive sentences. In the Discussion section, we will suggest other explanations for this finding.

Second, the model correctly reproduces the essential language differences in the effect of first noun animacy by demonstrating a significant interaction between Language and N1-Animacy. Moreover, the lower panel of Fig. 5 demonstrates that, in German, the benefit from animate first nouns is more pronounced in SVO sentences. This result matches the human data quite well.

Finally, any effects and interactions involving N2-Marking did not account for a significant proportion of variance in the network performance. This supports the experimental finding that strong cues provided early in the sentence result in fast, immediate responses which are not affected by any information provided at a later point unless they result in a complete revision of the initial interpretation. However, the network underestimated the magnitude of the reaction time increase that is required for this restructuring of the initial interpretation in sentences with unmarked first and nominative-marked second nouns.

Taken together, the network succeeded in capturing the essential language differences by exhibiting larger effects of first noun case-marking in Russian and larger effects of first noun animacy in German although it was unable to reproduce the differences between nominative and accusative-marking. This supports the assumption that the differential effects of nominative and accusative are unrelated to the distributional characteristics of case-markers in Russian and German.

Discussion

The recurrent cascaded backpropagation network matched the behavioural data obtained in the experiment closely in a variety of aspects which are summarised below. At the same time, there was one aspect in which the model failed to match up well with human reaction time data.

Availability effects. Both in the model and in the data, no significant language differences were found with respect to the final outcome of processing. The crosslinguistic differences related to the availability of cues manifest themselves in the time course of processing. Figure 7 illustrates how this effect comes about. While the asymptotic activation values arrived at in both languages are very similar, the time needed to reach the activation threshold may be different. These differences are larger at the beginning of the sentence suggesting that availability effects can best be detected in speeded tasks.

Case-marking effects. The model correctly simulates the fact that the processing benefit from case-marking is larger in Russian than in German. We will briefly describe the nature of this effect with respect to the patterns illustrated in Fig. 7. In Russian, threshold activation for sentences without case-markers is reached at the average after the third string (Fig. 7, first panel). The Russian network takes more time than the German network because, at the beginning of each pattern, the Russian network starts out with an activation value close to 0.5, as opposed to 0.9 in German. If the first noun is case-marked, the activation threshold can be reached after the first string, and the relative gain in time is quite substantial (Fig. 7, second and fourth panels). In the German network, the initial activation strongly favours the 'agent first' interpretation (around 0.9). Thus, in patterns with two unmarked nouns, the threshold can be reached at the second string (Fig. 7, upper panel), despite the absence of a reliable marker. Compared to this faster response the relative gain from case-marking on the first noun is smaller in German than in Russian. This

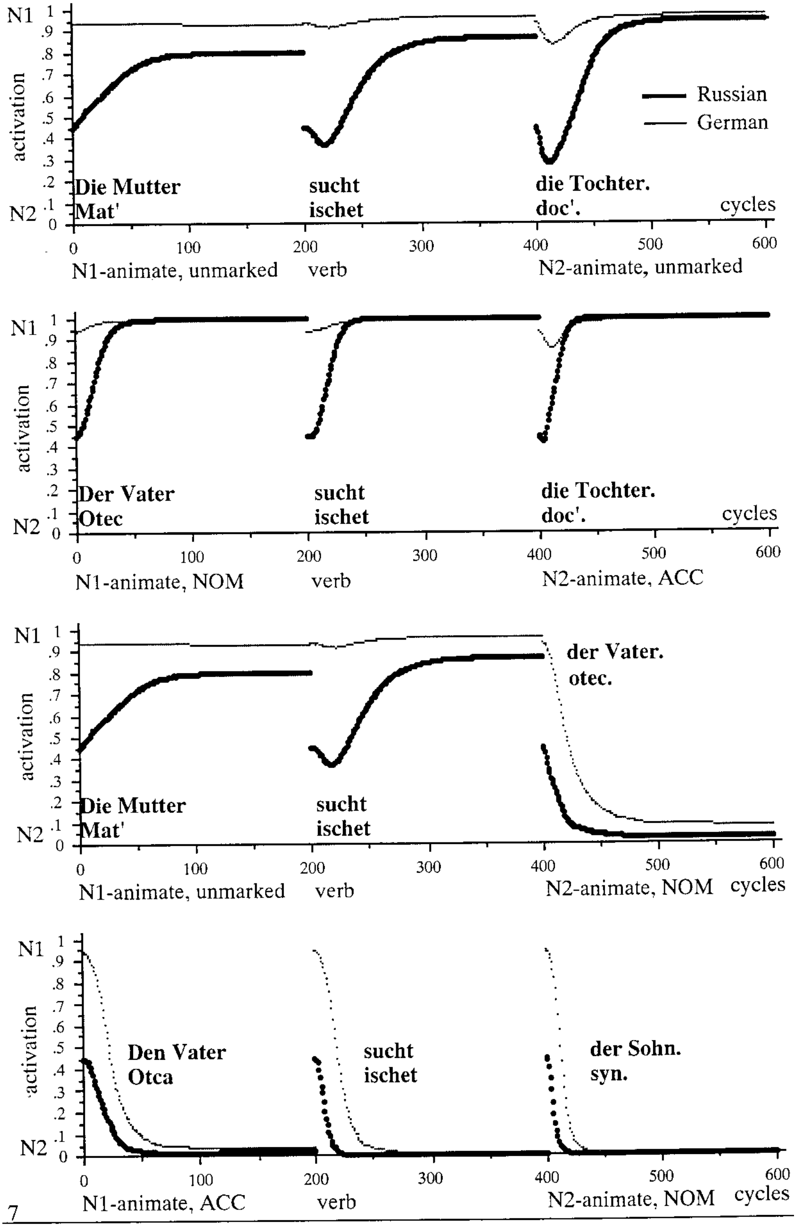

FIG. 7. Time course of activation build-up for selected patterns in the Russian (thicker line) and German (thinner line) networks.

analysis indicates that the larger benefit from case-marking in Russian is a relative one caused by different initial biases at the outset of processing. It should be noted, however, that the magnitude of this effect was greatly exaggerated in the model. This is most likely a consequence of the fact that the model was exposed to a very limited set of training patterns which coded only nominative and accusative-marking. An inclusion of a larger variety of sentence types containing all possible case-markers should diminish this exaggeration and lead to more realistic effect sizes. However, the main emphasis of this simulation was to replicate the language differences and, in this respect, it performed well.

Animacy effects. The network correctly captures the fact that animacy is a stronger cue in German than in Russian. This occurs not because animacy is more reliable or more readily available in German, but because case-marking is less available in German than in Russian. The simulation shows that the strength of any individual cue depends on the relative strength of other cues which, in turn, is determined by the distributional characteristics of the primary linguistic data.

On-line cue interaction. As in the human data, cues on the second noun did not lead to any additional benefit if they merely supported cues encountered earlier in the sentence. Thus, early cues appear to trigger immediate responses the network equivalent of which is a fast attainment of the criterion activation value. Only if later cues are in conflict with earlier ones, additional processing time for restructuring the initial interpretation is required. This aspect of the data was not matched by the network performance. However, a closer look at the network performance shows that this is a consequence of the way reaction times were estimated: even though the asymptotic activation value at 600 cycles was below criterion in patterns with unmarked first nouns and nominative-marked second nouns, the cut-off at 600 cycles leads to a systematic underestimation of any time required for additional processing.

Accusative superiority. The principal failure of the model was its inability to simulate the stronger effect of accusative-marking over nominative-marking of the first noun in the reaction times. We believe that the accusative superiority effect cannot be attributed to distributional characteristics of the nominative and the accusative in Russian and German, but, rather, is a result of the structure of the inflectional paradigms in the two languages. Each of the Russian and German inflections occurring in the nominative and the accusative case also serves as a marker for a variety of other cases dependent on the number and gender of the noun. For example, the German article *der* is not only associated with nominative in masculine nouns but also with genitive and dative in feminine nouns and with genitive in plural nouns. Similarly, the German article *den* is not only a marker for accusative in masculine nouns but also a marker for dative in plural nouns. The fact that *den* is associated with only two possible alternatives (accusative, dative) whereas *der* is associated with three alternatives (nominative, genitive, dative) suggests that the former is a less ambiguous case-marker than the latter. This problem of inflectional syncretism is even more severe in Russian where a very limited number of suffixes is distributed over a paradigm with 72 cells. The accusative superiority observed in the experiment might be a consequence of the fact that the inflections associated with the accusative are less syncretic, i.e. refer to a smaller number of cases across the language than the inflections associated with the nominative, even if they are perfectly reliable with respect to the individual noun. In order to capture this in a connectionist simulation, we will need to provide a richer representation of the complete inflectional paradigm. This is clearly a task for future modelling work.

GENERAL DISCUSSION

Two studies were designed to extend the Competition Model framework to account for crosslinguistic differences in on-line sentence processing. In particular, we were interested in assessing the possible role of cue availability as a determinant of on-line processing. The results of the experiment confirmed the existence of language differences in Russian and German on-line processing, despite the fact that both languages provide the same repertoire of reliable cues. We argue that the observed crosslinguistic differences in on-line processing must be attributed, generally speaking, to differences in cue availability. Moreover, the concepts of reliability and availability, taken alone, are not sufficient to predict more complex patterns of cue interactions. To account more satisfactorily for these effects, we constructed a recurrent cascaded backpropagation model that exhibited a good fit to the human behaviour observed in the experiment and was able to capture the essential language differences as well as the patterns of cue interactions.

The simulation provided a compelling illustration of the ways in which the effects of various cues unfold through real time. In the network, comprehension is modelled as incremental accumulation of evidence for competing alternative interpretations (e.g. ‘the first noun is the agent’ *vs* ‘the second noun is the agent’). Both the simulation and the data indicate that stronger cues permit to essentially ignore the presence of weaker cues whereas weaker cues allow other information to influence the timing and outcome of the sentence interpretation. The on-line cue interactions reflect relative differences in cue strength that can be traced to differences in cue availability stemming from neutralisations in the inflectional paradigms.

The development of these processing mechanisms can be described as a type of ‘linguistic tuning’ (Cuetos & Mitchell, 1988) in that the statistical properties of the language may tune the processing system to use the available cognitive resources in the most efficient way. The notions of cue integration and linguistic tuning are also compatible with both the interactive constraint-based view of sentence processing (MacDonald et al., 1994) and the referential-support approach (Altmann, 1988; Altmann et al., 1992; Altmann & Steedman, 1988). According to these models, semantic, referential, and contextual information as well as the statistical properties of the language all contribute to the selection of interpretation alternatives. What differentiates our approach from these complementary, related perspectives is our crosslinguistic methodological focus. While the majority of research in the domain of sentence processing has concentrated on the factors contributing to lexical and syntactic ambiguity resolution, we have focused on crosslinguistic differences in the processing of both non-ambiguous and ambiguous sentences in order to show how the statistical properties of the language modify the constraints operating in on-line processing. We have shown that ambiguity in the inflectional paradigms diminishes the availability of morphological cues which, in turn, decreases the on-line processing benefits associated with this cue. Use of this crosslinguistic methodology broadens the empirical basis for these theoretical claims beyond the scope of English to the realm of properties that may be true of language as a whole.

The fact that statistical properties of the language influence on-line processing is not immediately compatible with an approach that views language comprehension as the application of non-competitive rules or parsing principles (Frazier, 1987, 1989; Frazier & Fodor, 1978). However, the lack of research in the area of morphosyntactic processing within this approach makes it difficult to evaluate its specific predictions with respect to the effects of morphological cues. The fact that the orthogonal crossing of animacy and case-marking in the experiment revealed a stronger animacy benefit in German seems to argue against a view of morphosyntax as an encapsulated module insulated against semantic information (Bock, Loebell, & Morey, 1992; Fodor, 1983). Instead, the larger animacy effect in German suggests that language processing operates in a highly interactive fashion which permits the relatively weak evidence from the morphosyntactic domain to be supplemented by evidence from the semantic domain.

However, the experimental results revealed differences in the on-line effects of nominative-marking and accusative-marking that could not be captured by the simulation. The failure to account for the accusative superiority effect is most likely a consequence of the limited number of sentence types provided as input to the model and the fact that our coding of linguistic information represented nominative-marking and accusative-marking as dichotomous features thereby neglecting the specific inflectional changes associated with case. Future modelling attempts will have to include a sampling of the entire inflectional paradigm across a larger variety of sentence types in order to get a more precise picture of how human language processing adapts to the statistical properties of the language.

Taken together, this study has yielded five major findings, each of clear theoretical importance. First, we have shown that cue availability has a major on-line impact on sentence processing. Second, we have provided additional evidence for the nonmultiplicative effects of redundant cues for on-line processing. Third, we have shown that the absence of a single strong on-line cue in a language like German tends to provide room for the expansion of a secondary cue such as animacy. Fourth, we have shown that these on-line cue processing effects can be captured through a connectionist model that utilises a cascaded recurrent backpropagation architecture. Fifth, we have demonstrated crosslinguistic differences in the immediacy of processing of morphological cues such as case markers. We have shown that even two languages that have a parallel set of reliable morphological cues can differ in regard to on-line processing. Given this, we can expect even stronger crosslinguistic differences in on-line processing between languages that have markedly different cue repertoires, such as Navajo, Quechua, Arabic, or Georgian. We hope that these initial results further encourage other researchers to consider the use of systematic crosslinguistic comparisons as ways of understanding fundamental issues in language processing.

Manuscript received September 1996
Revised manuscript received January 1998

REFERENCES

Altmann, G.T. (1988). Ambiguity, parsing strategies, and computational models. *Language and Cognitive Processes*, *3*, 73–97.

Altmann, G.T., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191–238.

Altmann, G.T., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: eye movements in context. *Journal of Memory and Language*, *31*, 685–712.

Bates, E., MacWhinney, B., Caselli, C., Devescovi, A., Natale, F., & Venza, V. (1984). A crosslinguistic study of the development of sentence interpretation strategies. *Child Development*, *55*, 341–354.

Bates, E., McNew, S., MacWhinney, B., Devescovi, A., & Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition*, *11*, 245–299.

Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, *99*, 150–171.

Cohen, J., Dunbar, K., & McClelland, J.L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop-effect. *Psychological Review*, *97*, 332–361.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphical system for designing and controlling experiments in the Psychology laboratory using Macintosh computers. *Behavioral Research Methods, Instrumentation, and Computation*, *25*, 257–271.

Comrie, B., & Corbett, G. (Eds) (1992). *The Slavonic Languages*. London: Routledge.

Corrigan, R. (1986). The internal structure of English transitive sentences. *Memory and Cognition*, *14*, 420–431.

Corrigan, R. (1988). Who dun it? The influence of actor-patient animacy and type of verb in the making of causal attributions. *Journal of Memory and Language*, *27*, 447–465.

Cuetos, F., & Mitchell, D.C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition*, *30*, 73–105.

Elman, J. (1990. Finding structure in time. *Cognitive Science*, *14*, 179–212.

Fodor, J.A. (1983). *The Modularity of the Mind*. Cambridge, MA: MIT Press.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.) *Attention and performance XII* (pp. 601–681). Hove, UK: Lawrence Erlbaum Associates Ltd.

Frazier, L. (1989). Against lexical generation of syntax. In W. Marslen-Wilson (Ed.) *Lexical Representation and Process*. Cambridge, MA: MIT Press.

Frazier, L., & D'Arcais, G.B.F. (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language*, *28*, 331–344.

Frazier, L., & Fodor, J.D. (1978). The sausage-machine: A new two-stage parsing model. *Cognition*, *6*, 291–325.

Gernsbacher, M.A., Hargreaves, D.J., & Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language*, *28*, 735–755.

Hernandez, A., Bates, E., & Avila, L. (1994). On-line sentence interpretation in Spanish-English bilinguals: What does it mean to be "in between?". *Applied Psycholinguistics*, *15*, 417–446.

Kail, M. (1989). Cue validity, cue cost, and processing types in French sentence comprehension. In B. MacWhinney & E. Bates (Eds), *The Crosslinguistic Study of Language Processing* (77–117). New York: Cambridge University Press.

Kilborn, K. (1989). Sentence processing in a second language: The timing of transfer. *Language and Speech*, *32*, 1–23.

Li, P., Bates, E., & MacWhinney, B. (1993). Processing a language without inflections: A reaction time study of sentence interpretation in Chinese. *Journal of Memory and Language*, *32*, 169–192.

MacDonald, M.C., Pearlmutter, N.J. & Seidenberg, M.S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.

MacWhinney, B. (1977). Starting points. *Language*, *53*, 152–168.

MacWhinney, B. (1987). Toward a psycholinguistically plausible parser. In S. Thomason (Ed.) *Proceedings of the Eastern States Conference on Linguistics* (pp. 1–8). Ohio State University: Columbus. Ohio.

MacWhinney, B. (1995). The CHILDES project: Tools for analyzing talk (2nd ed.). Hillsdale, NJ: Erlbaum.

MacWhinney, B., & Bates, E. (Eds) (1989). *The Crosslinguistic Study of Sentence Processing*. New York: Cambridge University Press.

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, *23*, 127–150.
MacWhinney, B., Pléh, C., & Bates, E. (1985). The development of sentence interpretation in Hungarian. *Cognitive Psychology*, *17*, 178–209.
McClelland, J.L. (1979). On the time-relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, *86*, 287–330.
McDonald, J.L. (1986). The development of sentence comprehension strategies in English and Dutch. *Journal of Experimental Child Psychology*, *41*, 317–335.
McDonald, J.L., & MacWhinney, B. (1989). Maximum likelihood models for sentence processing research. In B. MacWhinney, & E. Bates (Eds), *The Crosslinguistic Study of Sentence Processing* (pp. 397–421). New York: Cambridge University Press.
McDonald, J.L., & MacWhinney, B.J. (1985). The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language*, *34*, 543–566.
Mimica, I., Sullivan, M., & Smith, S. (1994). An on-line study of sentence interpretation in native Croatian speakers. *Applied Psycholinguistics*, *15*, 237–261.
Mitchell, D.C. & Cuetos, F. (1991). The origins of parsing strategies. In C. Smith (Ed.) *Current Issues in Natural Language Processing* (pp. 1–12). Center of Cognitive Science, University of Austin, Texas.
Mitchell, D.C., Cuetos, F., & Zagar, D. (1990). Reading in different languages: Is there a universal mechanism for parsing sentences? In D.A. Balota, G.B. Flores d'Arcais, K. Rayner (Eds.) *Comprehension Processes in Reading* (pp. 285–302). Hillsdale, NJ: Lawrence Erlbaum Associates.
Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning internal representations by back propagation. In D. Rumelhart, & J. McClelland (Eds), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (pp. 318–362). Cambridge, MA: MIT Press.
Snodgrass, J.G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174–215.
Sokolov, J. (1989). The development of role assignment in Hebrew. In B. MacWhinney, & E. Bates (Eds), *The Crosslinguistic Study of Sentence Processing* (pp. 118–157). New York: Cambridge University Press.
Taraban, R. & Roark, B. (1996). Competition in learning language-based categories. *Applied Psycholinguistics*, *17*, 125–148.
Trueswell, J.C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language*, *35*, 566–585.
Trueswell, J.C., Tanenhaus, M.K., & Garnsey, S.M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*, 285–318.
Vigliocco, G., Butterworth, B., & Semenza, C. (1994). Constructing subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language*, *34*, 186–215.
Winer, B.J. (1971). *Statistical Principles in Experimental Design*. New York: McGraw Hill.
Zubin, D. (1979). Discourse function of morphology: The focus system in German. In T. Givón (Ed.), *Syntax and Semantics: Discourse and Syntax* (Vol. 12, pp. 469–504). New York: Academic Press.
Zubin, D.A. (1977). The semantic basis of case alternation in German. In R.W. Fasold, & R.W. Shuy (Eds), *Studies in Language Variation: Semantics, Syntax, Phonology, Pragmatics, Social Situations, Ethnographic Approaches* (pp. 88–89). Washington, DC: Georgetown University Press.

APPENDIX 1

Source references for the corpus analysis in Russian and German

German

Danella, U. (1989). *Die Unbesiegte*, München: Wilhelm Heyne. Pp. 169–173, 459–463.

Die Kinder- und Hausmärchen der Brüder Grimm (1967). Der Kinderbuchverlag: Berlin. Pp. 136–139.

Rickenhaus, Fr. (1993). Wieviel Programme erträgt ein Mensch? *Die Zeit*, *48* (Nov. 23), 13–15.

Schmidt-Häuer, Chr. (1993). Die Diktatur des Demokraten. *Die Zeit*, *47* (Nov. 19).

Simmel, J.M. (1988). *Die im Dunkeln sieht man nicht*. München: Knauer. Pp. 75–80, 388–392.

Russian

Antonov, S. (1976). Povesti i passkazy, Moskva. *Molodaja Gvardija*, 86–90, 210–214.

Bitov, A. (1993). Ozhidanije obes”jan (Fragment romana). *Moskovskije Novosti*, *45*, Nov. 7.

Gorbunov, A. & Kokesnikov, A. (1993). Shturm “Belogo Doma”: “Al’fa” ne khotela krovi. *Moskovskije Novosti*, *45*, Nov. 7.

Navrozov, L. (1993). U nas etogo citat’ ne budut. *Literaturnaja gazeta*, *47*, Nov. 24.

Russkije Skazki (1991). Moskva, *Sovetskaja Rossija*, 34–44.

Strokan’, S. (1993). Yadernaja Ukraina—bol’shoj blef ili bol’shaja ugroza. *Moskovskije Novosti*, *45*, Nov. 7.

Vainer, A.A. & Vainer, G.A. (1985). *Era miloserdija.* Kishin’ov, Lumina, 115–119, 279–283, 377–381.

Vassil’ev, G. (1993). Somnenie est’ kompliment iskusstvu, *Moskovskije Novosti*, *45*, Nov. 11.

APPENDIX 2

Design and example sentences used in the picture choice task

Cell no.	*N1-Animacy*	*N2-Animacy*	*Configuration*	*N1-Marking*	*N2-Marking*	*German*	*Russian*
1	anim	anim	SVO	unmark	unmark	Die Mutter sucht die Tochter.	Mat' iščet doč'.
2					marked	Die Mutter sucht den Sohn.	Mat' iščet syna.
3				marked	unmark	Der Vater sucht die Tochter.	Otec iščet doč'
4					marked	Der Vater sucht den Sohn.	Otec iščet syna.
5			OVS	unmark	unmark	Die Tochter sucht die Mutter.	Doč' iščet mat'.
6					marked	Die Tochter sucht der Vater.	Doč' iščet otec.
7				marked	unmark	Den Sohn sucht die Mutter.	Syna iščet mat'.
8					marked	Den Sohn sucht der Vater.	Syna iščet otec.
9		inanim	SVO	unmark	unmark	Die Mutter sucht die Torte.	Mat' iščet tort.
10					marked	Die Mutter sucht den Teller.	Mat' iščet tarelku.
11				marked	unmark	Der Vater sucht die Torte.	Otec iščet tort.
12					marked	Der Vater sucht den Teller.	Otec iščet tarelku.
13			OVS	unmark	unmark	Die Tochter sucht die Blume.	Doč' iščet cvetok.
14					marked	Die Tochter sucht der Löffel.	Doč' iščet ložka.
15				marked	unmark	Den Sohn sucht die Blume.	Syna iščet cvetok.
16					marked	Den Sohn sucht der Löffel.	Syna iščet ložka.
17	inanim	anim	SVO	unmark	unmark	Die Blume sucht die Tochter.	Cvetok iščet doč'.
18					marked	Die Blume sucht den Sohn.	Cvetok iščet syna.
19				marked	unmark	Der Löffel sucht die Tochter.	Ložka iščet doč'
20					marked	Der Löffel sucht den Sohn.	Ložka iščet syna.
21			OVS	unmark	unmark	Die Torte sucht die Mutter.	Tort iščet mat'.
22					marked	Die Torte sucht der Vater.	Tort iščet otec.
23				marked	unmark	Den Teller sucht die Mutter.	Tarelku iščet mat'.
24					marked	Den Teller sucht der Vater.	Tarelku iščet otec.

APPENDIX 2 *(Contd)*

Design and example sentences used in the picture choice task

Cell no.	*N1-Animacy*	*N2-Animacy*	*Configuration*	*N1-Marking*	*N2-Marking*	*German*	*Russian*
25		inanim	SVO	unmark	unmark	Die Blume sucht die Torte.	Cvetok iščet tort.
26					marked	Die Blume sucht den Teller.	Cvetok iščet tarelku.
27				marked	unmark	Der Löffel sucht die Torte.	Ložka iščet tort.
28					marked	Der Löffel sucht den Teller.	Ložka iščet tarelku.
29			OVS	unmark	unmark	Die Torte sucht die Blume.	Tort iščet cvetok.
30					marked	Die Torte sucht der Löffel.	Tort iščet ložka.
31				marked	unmark	Den Teller sucht die Blume.	Tarelku iščet cvetok.
32					marked	Den Teller sucht der Löffel.	Tarelku iščet ložka

APPENDIX 3

F-*values and effect sizes in the first noun choices (human data and network)*

	Proportion N1 choice		*% experimental variance*		
Effect	$F_1(1,26)$	$F_2(1,193)$	*By subjects*	*By items*	*Network*
N1-Animacy	18.5***	46.4***	1.1	1.6	1.6
N2-Animacy	9.6**	42.5***		1.4	
Configuration	864.9***	1784***	61.2	60.6	60.1
N1-Marking	296.5***	225.8***	8.1	7.7	2.9
N2-Marking	52.9***	21.2***	1.3	1.0	3.6
Configuration × N1-Marking	619.8***	440.2***	14.4	14.9	13.5
Configuration × N2-Marking	138.5***	142.6***	4.8	4.8	4.8
N1-Animacy × N2-Animacy	7.4*	18.8***			
N1-Animacy × N1-Marking	23.8***	22.4***			1.8
N1-Animacy × N2-Marking	16.8***	10.1***			1.8
N1-Marking × N2-Marking	55.3***	27.2***	1.3		2.8
Configuration × N1-Marking × N2-Marking	104.4***	109.9***	4.2	3.7	3.6

Effect sizes (in percentage of experimental variance) are given only if the effect or interaction accounted for more than 1% of variance. *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

APPENDIX 4

F-values and effect sizes in the reaction times for the human data and the number of cycles until criterion in the connectionist network

Effect	Reaction times adjusted to the end of 1st noun		Reaction times adjusted to the end of verb		Reaction times adjusted to the end of 2nd noun		cycles % var network
	$F_1(1,26)$	$F_2(1,193)$	$F_1(1,26)$	$F_2(1,193)$	$F_1(1,26)$	$F_2(1,193)$	
N1-Animacy		24.8*** (6.0)		21.5*** (6.0)		4.3* (11.0)	4.8
N2-Animacy	4.1*	2.3	4.7*	2.0	8.7** (1.6)	11.3** (2.5)	
N1-Marking	53.3*** (45.6)	188.1*** (45.2)	50.1*** (42.1)	167.3*** (46.5)	49.9*** (37.3)	167.5*** (36.8)	81.4
N2-Marking	21.8** (6.6)	24.9*** (6.0)	20.5*** (6.1)	22.1*** (6.1)	8.3** (2.2)	8.0** (1.8)	
Language × N1-Animacy	10.2*** (2.3)	3.1	12.3** (2.4)	2.9	13.5** (2.3)	7.2** (1.6)	2.3
Language × N1-Marking	3.8	10.5** (2.5)	3.3	9.3** (2.6)	3.2	9.3** (2.0)	3.0
Configuration × N1-Marking	25.1*** (13.7)	51.3*** (12.3)	25.1*** (13.0)	45.7*** (12.7)	24.6*** (11.2)	46.1*** (10.1)	
Configuration × N2-Marking	13.0*** (1.8)	8.9** (2.1)	10.4** (1.5)	7.9* (2.2)	12.6** (1.6)	9.1** (2.0)	
N1-Marking × N2-Marking	18.4*** (4.2)	13.8*** (3.3)	18.1*** (3.9)	12.3** (3.4)	19.5*** (3.5)	12.1** (2.7)	
Language × Configuration × N1-Animacy	3.0	8.2** (2.9)	3.5	7.6** (2.1)	3.5	3.9* (1.0)	
Language × Configuration × N1-Marking	4.4** (2.4)	9.2** (2.2)	4.5*	8.2** (2.3)	4.4* (2.0)	8.3** (1.8)	
Configuration × N1-Animacy × N2-Marking	4.1	3.6*	4.4*	3.0*	5.1*	4.7* (1.0)	
Configuration × N1-Marking × N2-Marking	13.5** (1.7)	5.5* (1.3)	12.6** (1.5)	4.9* (1.3)	14.2** (1.5)	4.9* (1.3)	

Effect sizes (in percentage of experimental variance) are given only if the effect or interaction accounted for more than 1% of variance. *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$, no asterisk: $0.05 < P < 0.1$.