Visuo-Linguistic Interaction in Complex Environments: Understanding Navigation Instructions in Driving

Hendrik Koesling (hendrik@cs.nuim.ie)
Department of Computer Science, National University of Ireland
Maynooth, Co. Kildare, Ireland

Ronan G. Reilly (ronan@cs.nuim.ie)
Department of Computer Science, National University of Ireland
Maynooth, Co. Kildare, Ireland

Abstract
Driver-assistant devices such as navigation systems are now a common feature in many cars. Processing navigation instructions, however, is a challenging task and has to be accomplished in parallel to the actual driving task. In order to improve driving safety, we must optimise the cognitive workload spent on each task. With the cognitive workload distribution constantly changing between the two concurrent tasks in dynamic driving environments, navigation messages should be adapted according to the current traffic situation. We investigated how drivers integrate verbal and visual information while viewing animated traffic scenes and responding to spoken navigation instructions. Linguistic and visual factors showed significant effects on instruction comprehension. Oculomotor data and the distribution of visual attention reflected this visuo-linguistic interaction, indicating how drivers process navigation messages in dynamic situations. Results show, for example, that vehicular traffic and road furniture were mostly attended to before and after messages were issued. As soon as messages were issued, the focus of attention shifted to areas that were verbally referred to. We found a general pattern of serial mapping of gaze points onto verbally referred-to scene objects. This pattern, however, often exposed characteristic “interruptions” when additional, relevant scene information was processed. Furthermore, complex instructions in crowded traffic scenes or redundant verbal information compromised the parallel processing and integration of verbal and visual information. Apparently, a tight synchronisation of auditory and visual input is required for the accurate comprehension of an instruction. If this process is disturbed, message comprehension in complex traffic environments deteriorates and may lead to dangerous driving manoeuvres.

Introduction
Numerous studies have investigated human-machine interfaces with a particular focus on usability. In order to achieve “natural” interaction with artificial systems and to generate adequate machine response, it is necessary to understand characteristic human behaviour in specific situations. In this context, language and vision are two of the most relevant communication modalities and sources of information. For many such interfaces it is sufficient merely to function in static environments. Some important ones, however, operate in highly dynamic environments. For the latter type it is essential to adapt to changing situations. In other words, the system itself has to be dynamic.

A prototypical example of human-machine interaction in dynamic environments is an in-car driver-assistant system that communicates with the driver (Barfield & Dingus, 1998; Gale,
Brown, Haslegrave & Taylor, 1998; Noy, 1997). One of the developments in this area is in-car navigation systems. While travelling, they issue messages that direct the driver to a specified destination. It is obvious that in situations such as driving, many cognitively challenging tasks have to be accomplished in parallel. Navigation messages must therefore be generated carefully in order to ensure easy and correct comprehension: The less mental effort spent on such secondary tasks and their cognitive processing, the more the driver can concentrate on the primary task, the actual driving. This is highly relevant for driving safety (e.g. Recarte & Nunes, 2000). Due to the dynamics of the environment during driving, the distribution of the cognitive workload between the concurrent tasks constantly changes. It would thus be desirable to adapt navigation messages according to the current traffic situation. Therefore, the linguistic factors that determine navigation messages should be adjusted in response to changes in the visual factors that characterise traffic scenes. This could optimise the processing of concurrent tasks and improve driving safety.

A literature review produced a rather ambivalent image of the current state of research with respect to the processing of messages issued by driver-assistant systems and with respect to (non-language related) design aspects of such systems. On the one hand, numerous studies address the processing of spoken language in general – for an overview see, for example, Gernsbacher (1994), Rickheit (1999), Rickheit, Sichelschmidt and Strohner (2002) and Sucharowski (1996) – and that of route descriptions and automotive action directions in particular (e.g. Freska, Habel & Wender, 1998; Graham & Mitchell, 1997; Kimura, Marunaka & Sugiuira, 1997). On the other hand, several researchers examine the practical use of driver-assistant systems and focussed on usability and ergonomic factors (e.g. Dingus, Hulse & Barfield, 1998; Ross, Vaughan & Nicolle, 1997; Spyridakis, Miller & Barfield, 1998). These latter investigations, however, often lack the appropriate cognitive “grounding”. Only very few investigations exist that take into account aspects of both approaches (e.g. Liu, 1998). These successfully integrate the display of realistic in-car and driving environments as well as the fundamental requirements of language processing from the viewpoint of cognitive linguistics and cognitive psychology.

In an attempt to bridge the gap between the two approaches sketched above and inspired by our previous research (Rickheit, Sichelschmidt, Koesling & Barattelli, 1999), we investigated the interaction between language and visual attention during the presentation of spoken navigation messages in dynamic, animated driving scenarios. Oculomotor data collected in an eye-movement study provided detailed information regarding the spatio-temporal distribution of visual attention (e.g. Pomplun, 1998; Pomplun, Ritter & Velichkovsky, 1996), reflecting visuo-linguistic interaction during driving. This allowed us to draw conclusions about how drivers process navigation messages in dynamic situations, depending on linguistic and visual factors as well as the timing of navigation messages.

We specifically addressed the following research questions and hypotheses: How does the distribution of attention change between the different phases of driving while following navigation instructions, i.e. before, during and after the presentation of a navigation message? Can we – analogous to the incremental generation of representations from spoken language in static scenes (e.g. Eberhard, Spivey-Knowlton, Sedivy & Tanenhaus, 1995; Alloppenna, Magnuson & Tanenhaus, 1998; Spivey, Tanenhaus, Eberhard & Sedivy, 2002) – serially map gaze trajectories onto navigation instructions in complex, dynamic situations such as driving? Does a correlation exist between possible disturbances of this mapping and a deterioration of driver responses? Finally, which factors regarding the properties of the navigation instruction, the current traffic situation and their interaction (e.g. timing of messages) particularly enhance or compromise message understanding? The findings could help to formulate guidelines for the use of specific instruction characteristics, depending on the scene content. Navigation systems could thus dynamically adapt to the current traffic situation and improve driving safety by optimising the cognitive workload of the driver.
Experiments

We designed a set of four experiments that accounted for different experimental conditions regarding the relevant properties of navigation instructions (e.g. message complexity and verb position) and dynamic traffic scenes (e.g. scene complexity) and their interaction (e.g. message timing). In the principal experimental setting, subjects watched animated video sequences of typical road scenes. Spoken navigation-system messages were issued at various stages during the video presentation and the participants’ responses were monitored in order to assess the comprehension of the spoken messages and the implementation of the instructions – a classical example of a dual task with stimulation of both the visual and the auditory channels (Daneman & Carpenter, 1980; Liu, 2001; Wickens, Gordon & Liu, 1998).

Method

In total, 80 subjects took part in four experiments (20 subjects per experiment); 50% were female, 50% male. All subjects were experienced drivers and held their driving licence for at least 6 years ($\mu_{\text{license}} = 9.4; \sigma_{\text{license}} = 2.3$).

As “directional” instructions (e.g., “Turn right after the bridge.”) are the most common messages issued by navigation systems, we focussed on this type of instructions. In order to systematically investigate the visuo-linguistic interaction during the processing of such messages while driving, we defined a set of road scenes together with a set of spoken instructions to form the experimental stimuli. We varied those factors that best reflect typical road scenes (visual factors) and those that must be considered essential for message comprehension in time-critical situations (linguistic factors).

With respect to the visual factors, we introduced the construct of “scene complexity”. In “simple” scenes only two vehicles were visible (including the subject’s own vehicle), and the turning alternative that the spoken instruction referred to was unambiguous. In “complex” scenes, four vehicles were visible and several turning alternatives had to be scanned to find the one that was referred to in the spoken instruction.

With respect to the linguistic factors, we systematically varied the formal, syntactic and semantic structures of spoken instructions. We chose message length (formal structure) with instructions either being short (2 words), intermediate (5 words) or long (7 or 8 words). This categorisation coincided with three levels of a specific semantic structure property, namely the number of propositions (2, 4, or 6). The joint factor “message complexity” thus emerged, determined by the instructions’ formal and semantic structures. Message complexity levels were “simple”, “intermediate” and “complex”. In order to account for the syntactic message structure, the factor “verb position” was varied. The verb was placed either at the start or the end of an instruction.

“Message timing” was also varied (“early” or “late”, message ends 3 or 1 seconds before a turning point is reached, respectively), manifesting the interaction between visual and linguistic properties.

This design led to a series of four experiments. The factors message timing and verb position varied between experiments (between-subjects factors), whereas the factors scene and message complexity varied within experiments (within-subject factors). Each of the within-subject factor combinations (2x3) was repeated four times so that all subjects had to complete 24 trials. Table 1 illustrates how conditions were varied in the experiment series.

Stimuli consisted of pre-recorded video sequences of abstract road scenes viewed from a bird’s eye perspective. Although it might not resemble the driver’s normal view too closely, the bird’s eye perspective appeared to be more appropriate for the investigation. First, the whole scene remains visible at all times. This guarantees the persistence of stimulus factors throughout a trial, for example the number of vehicles visible. Furthermore, reducing stimuli to abstract road scenes with only essential road furniture minimises distraction and should yield more reliable statistical results with respect to the influence of visual factors on eye-movement parameters in particular. It also helps to focus the subjects’ attention on the actual task. The video sequences were overlaid with the spoken navigation instructions, yielding the
experimental stimuli. Figure 1 shows (greyscale) still images of typical animated stimuli used in Experiments 1 (left) and 3 (right), showing the track layout with road markings, traffic signs and several vehicles. In the experiments, we used colour stimuli.

Stimuli were presented on a 19” colour computer monitor with a CRT screen; subjects were seated at a distance of 60 cm from the display. Eye movements were recorded with an SR Research EyeLink II eye-tracker. All subjects were tested individually. Subjects had to complete 24 experimental trials.

A fixation marker appeared at the centre of the screen for 300 ms at the beginning of each trial. Immediately after, the pre-recorded video sequences started. Subjects had to press either the left or the right response button corresponding to the direction specified in the spoken instruction issued during the video display. Subjects should press the response button at the time when their vehicle reached the junction referred to in the instruction: “left/right” response. The screen was blanked and subjects had to state if it was safe to follow the instruction at the time of arrival at the junction, taking into account the road furniture and the other vehicles’ routes: “safe/unsafe” response. Turning right at a “Yield” sign while another vehicle approaches, for example, would require an “unsafe” response. Prior to the experiments, an expert judged the stimuli and classified situations as “safe” or “unsafe”. Numbers of “safe” and “unsafe” situations were balanced in the experiments.

The overall lengths of the video sequences – until a response button press would normally be required – varied between 5 and 10 seconds, depending on the visual and linguistic factors. In order to ensure equal orientation periods in all trials, each video started 2 seconds before

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Visual complexity</th>
<th>Linguistic complexity</th>
<th>Verb position</th>
<th>Message timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>All</td>
<td>All</td>
<td>Start</td>
<td>Early</td>
</tr>
<tr>
<td>2</td>
<td>All</td>
<td>All</td>
<td>End</td>
<td>Early</td>
</tr>
<tr>
<td>3</td>
<td>All</td>
<td>All</td>
<td>Start</td>
<td>Late</td>
</tr>
<tr>
<td>4</td>
<td>All</td>
<td>All</td>
<td>End</td>
<td>Late</td>
</tr>
</tbody>
</table>

Table 1: Overview of experimental design

Figure 1: Still images of animated experimental stimuli. The subject’s vehicle is white, the other vehicles are black. The arrow below the instruction transcript indicates the position of the word currently uttered. Left: Simple scene, simple instruction, verb at beginning of instruction, message issued early. Right: Complex scene, complex instruction, verb at beginning of instruction, message issued late.
the spoken instruction commenced. The subjects’ vehicle always started travelling on the same road at the bottom centre of the screen.

Results and Discussion
We analysed the influence of the visual and linguistic factors and the message timing on traditional psychophysical measures and on various eye-movement parameters. We will here only discuss some of the more relevant measures: response correctness, number of fixations and gaze duration. More specifically, we considered the ratios of numbers of fixations (rNF)\(^1\) and gaze durations (rGD)\(^2\) in designated regions for particular time intervals. This allowed us to obtain information about the distribution of visual attention during the scene display. We divided the stimulus display into four spatial regions: areas of verbal reference which were explicitly referred to in the navigation instructions (label: “instruc”), areas of interest (“interest”) with respect to the implementation of the instruction (e.g. road furniture), traffic areas, i.e. the locations of vehicles (“vehicles”), and areas irrelevant for the task (“irrelev”). As these areas varied in size, rNF and rGD were normalised.

We also chose the following three time intervals: before (I), during (II) and after the issuing of a spoken instruction (III). Analyses of variance were computed for all variables, Scheffé post-hoc tests determined which means differed for factors with more than two levels. For improved readability, only the most relevant statistical figures are quoted in the text.

In all four experiments, more than 97% of all “left/right” responses were correct, independent of all factors. The high ratio of correct responses was not unexpected as subjects had to memorise only a single word from the instruction. In case of the “safe/unsafe” response, correctness reached an overall average of 81.2%. However, significant differences exist for within-subject factors as well as for between-subjects factors: in all experiments simple scenes were significantly more correctly judged than complex ones: 87.6% vs. 82.2% (F(1;76) = 57.33; p < 0.001) and simple messages (85.9%) were significantly more correctly judged than intermediate (82.6%) and complex messages (80.4%) (F(2;152) = 61.96; p < 0.001). The comparison between experiments showed that “safe/unsafe” response correctness was significantly higher when messages were issued early: 85.1% vs. 80.1% for late messages (F(1;76) = 77.02; p < 0.001). The most notable interaction between within-subject factors and between-subjects factors demonstrated a significantly higher error rate for complex messages, irrespective of scene complexity, when messages were issued late and the verb position was at the end of an instruction. Subjects apparently have no problem correctly processing messages that are short and contain only few propositions in complex environments – even when the verb, and consequently the direction as well, is at the end of an instruction.

The analysis of the eye-movement parameters and the distribution of attention, dynamically changing with the scene, could now offer some explanations for the misjudgements of the traffic situations. More specifically, this could enable us to understand the nature of the interaction between visual scene elements and spoken instructions and how it can facilitate or compromise “safe” behaviour in automotive environments.

Taking into account the rNFs and rGDs, distinctive fixation patterns and distributions of attention exist for the specified time intervals. Before a message was issued, traffic was mainly attended to: rGD\(_\text{vehicles} = 0.72\). In contrast, roads and road signs are only sparsely fixated: rGD\(_\text{interest} = 0.25\) (F(1;76) = 13.12; p < 0.001). This pattern was found in all four experiments and did not vary significantly with scene complexity. A more detailed investigation of the distribution of attention within this time interval revealed, however, that subjects often briefly scanned the stimulus at the very beginning of each trial (2-3 fixations, saccade lengths approx. 1/3 of the screen width/height). Although this fixation pattern was not consistent in all trials and subjects, it suggests the generation of a coarse scene overview. In order to access this assumption, further investigation into the detailed processes in this time interval is required.

---

\(^1\) Number of fixations in specific region relative to all fixations.

\(^2\) Gaze duration in specific region relative to overall gaze duration.
When spoken messages were issued (time interval II), visual attention generally shifted to areas that were verbally referred to (rGD instruc = 0.61) and that were of interest for task completion (rGD interest = 0.21). These include road furniture such as road signs and line markings on the roads, but also the junctions on the way to the “target” junction, i.e. where turning off is required. Although traffic movements must be considered important during this period also, only a few fixations landed on vehicles while the spoken message was issued (rGD vehicles = 0.15). This indicates that the visual focus was mainly guided by the verbal instruction, taking into account additional, task relevant information from the scene as well. 

As in all experiments for all factor combinations, rGD irrelev and rNF irrelev were below 0.05. rGDs differed significantly from each other (F (3;228) = 15.33; p < 0.001). Furthermore, rGDs differed with the within-subject and the between-subjects factors. Figure 2 illustrates the rNFs and rGDs as a function of stimulus area and message issuing time.

As data has already suggested, traffic situations could be more correctly judged (safe/unsafe response) when messages were issued early. This was in particular true for complex messages. However, when a complex message was issued early and the verb was at the beginning of the instruction, some confusion could arise and lead to apparently inefficient visual message processing. Eye-movement patterns of a number of subjects during the processing of an instruction such as “Turn left at second junction.” in Experiment 1 showed saccades to the first road on the left immediately after the direction (“left”) was issued. This did not occur in Experiment 2 where no such ambiguity exists in “At second junction left turn.”.

Redundant parts of messages were similarly critical, for example when subjects were instructed to “Turn left into minor road after bridge.” while only one possibility to turn left exists. As subjects sequentially mapped messages onto the visual reference points in the scene, the processing of redundant information distracted them from shifting attention to other task-relevant areas, for example, vehicular traffic or the designated areas of interest. This is of particular relevance in complex scenes with much traffic and a complex road layout. Redundant verbal information should thus be avoided in complex scenarios. In “quiet” road situations, however, it may be used as additional “backup” information.

Not surprisingly, visual attention was not entirely taken off vehicular movements while a spoken instruction was issued. However, the decrease in rNF vehicles and rGD vehicles in comparison to the previous time interval was much steeper than that in rNF interest and rGD interest, i.e. in areas that were of interest for the task, but not verbally referred to in the instruction. At which stage of the verbal instruction these interruptions occurred and what might have caused them still needs to be clarified. Preliminary results suggest that the guidance of visual attention by the verbal instruction sometimes paused when other vehicles were approaching the subjects’ vehicle. In such cases, several alternating saccades between the two vehicles occurred and each vehicle was pursued for a short time after each saccade.
probably to “clear” the traffic situation. This mostly happened in complex scenes or when complex messages were issued. In such cases usually more vehicles were in the scene, increasing the proximity of vehicles. Due to the increased length of the message, the other vehicles were also more likely to make progress towards the subjects’ vehicle.

When the “clearing” was completed, subjects attempted to resume processing of the message. They regularly skipped the missed part of the instruction and continued visual processing at the point that related to the current verbalisation. Data indicates that subjects did not shift attention to locations that were referred to in the skipped part of the instruction. This observation could offer an explanation for the findings that more complex messages and more

![Sample gaze trajectories. Left column: Simple scene, intermediate instruction complexity, verb at end of instruction, message issued early. Instruction: “At second junction left turn.”. Right column: Complex scene, intermediate instruction complexity, verb at beginning of instruction, message issued late. Instruction: “Turn left after pedestrian crossing.”. Trajectories are shown separately for the time intervals before (top), during (center) and after the spoken instruction is issued (bottom). Solid lines mark vehicles’ routes during each time interval.](image-url)
complex scenes resulted in lower response correctness. We might conclude that the perception and processing of the message is greatly compromised when visual attention moves away from the items currently referred to by the instruction. This would indeed suggest a strong visuo-linguistic interaction between message perception (auditory), message processing (linguistic) and referencing (visual) as well as message understanding (cognitive). Furthermore, it appears that verbal and visual information cannot be processed in parallel when an asynchrony exists between the current verbal (spoken) input and the focus of visual attention.

For some factor combinations, however, the “opposite” happens. The focus of visual attention does not have to “catch up” with the instruction, but instead is ahead of it, scanning a scene “in anticipation”. This often happened in simple scenes when the verb was at the end of the instruction. In such cases, the target junction was specified at the beginning of the instruction and the traffic situation was clear. This allowed for the checking of turning alternatives (left or right) already before they were verbalised. Also, in simple scenes with only one possible turn off point, the focus of attention regularly moved away from areas that were referred to in the instruction. Eye movements almost exclusively followed traffic movements instead and also accessed relevant road furniture on the way to the target junction. As the turn-off point was clear at an early stage, there was apparently no further need for a close instruction-eye gaze coordination.

The comparison of eye-movement data between experiments yielded further interesting results and conclusions with respect to the timing of messages. This is of particular relevance in the last phase (III) of the stimulus presentation after the verbal instruction was completed. Subjects then had either 3 or only 1 second left to respond (for early or late messages, respectively). If messages were issued early (Experiments 1 and 2), rGDs (and rNFs) for the vehicular traffic and area-of-interest regions increased again to rGD_{vehicles} = 0.32 and rGD_{interest} = 0.45, rGD_{instruc} dropped to 0.23 (F(3;228) = 21.47; p < 0.001). These values were also significantly different from those in the previous phase (F(1;76) = 18.21; p < 0.001). When the focus of attention moved to vehicular traffic, significantly more often the vehicles closest to the subjects’ vehicle and those between the current position of the subjects’ vehicle and the target junction were fixated and, for approx. 400 ms on average, pursued (smooth pursuit eye movements). Only when scene complexity was low, vehicles that were not within the areas of interest were briefly fixated as well. Figure 3 shows two sample gaze trajectories, separated for the three time intervals, i.e. before, during and after a spoken instruction was issued.

In Experiments 3 and 4, where messages are issued late, the distribution of attention suggests that traffic was more or less ignored. Whereas vehicles irrelevant for the scene were hardly fixated, even those of relevance for the task attracted only little direct visual attention. This is visible in significantly reduced rNF_{vehicles} and rGD_{vehicles} (0.13) in comparison to the respective figures in Experiments 1 and 2. However, it is also possible that due to the restricted time left to complete the task, subjects relied on peripherally acquired visual information rather than shifting attention. This would not be reflected in the eye-movements parameters. We must in fact assume that peripheral information plays some role here. A number of studies have shown that the direction and velocity of moving stimuli can be quite accurately assessed in the peripheral field of view. Subjects might thus have applied an efficient strategy under time-critical conditions such as those in the Experiments 3 and 4, guiding overt visual attention to those areas that most required foveal processing. This assumption appears to be supported by rNF_{interest} and rGD_{interest} rising to significantly higher ratios (rGD_{interest} = 0.63) than in Experiments 1 and 2 (F(1;76) = 23.16; p < 0.001 for rGD_{interest}). Road furniture and relevant intersections on the way to the target junction were located in the respective areas and must be considered in order to complete the task. The assessment of the information present at these locations such as details of road signs or line markings apparently could not be accomplished that easily in the peripheral field of view. Furthermore, areas of verbal reference were still being fixated with rGD_{instruc} = 0.21.
General Discussion

In summary, the analysis of eye-movement parameters demonstrated that the distribution of attention changed significantly during scene presentation. While the vehicular traffic was mostly attended to before a spoken navigation message was issued, the focus of attention shifted to areas that were verbally referred to while the message was issued. After the message was completed, main attention shifted back to traffic observation and to those areas that were of particular interest with regard to the implementation of the instruction given in the navigation message. Although this principal pattern was found in most trials, significant differences exist, depending on visual and linguistic factors and on the timing of messages.

With regard to message comprehension in dynamic automotive environments, the analysis of eye movements revealed how visuo-linguistic interaction affected visual attention. It appears that the visual focus is indeed guided by the verbal instruction, very similar to the incremental generation of representations from spoken language in static scenes (e.g. Eberhard, Spivey-Knowlton, Sedivy & Tanenhaus, 1995; Allopenna, Magnuson & Tanenhaus, 1998; Spivey, Tanenhaus, Eberhard & Sedivy, 2002). However, additional scene information is also taken into account. The amount of additional information that is also processed is largely influenced by the complexity of the scene and the message timing. For complex scenes and late messages, in particular when complex messages are processed, this additional information can be critical. If the sequential processing of information directly referred to in the message was interrupted for too long, and visual attention was directed to locations of interest that were not referred to in the message, the resumption of message processing caused problems, leading to an increase in error rates for message comprehension. In addition, ambiguous instructions, even when contained in normally correctly processed messages, or redundant information, seemed to interrupt the processing of messages. The parallel processing of verbal and visual information and their integration that allowed for the accurate comprehension of an instruction and the correct implementation of the required action would then be disturbed. The attempt to synchronise these two streams of information from the auditory and visual input channels largely compromised message understanding in such dynamic and often time-critical environments.

We are currently working on a computational model that appropriately simulates the serial mapping of gaze trajectories onto navigation instructions and also takes into account the effects of processing additional scene information in dynamically changing environments. We are also experimenting with physical driving environments in order to evaluate the ecological validity of the data obtained in simulated driving. The fact, that subjects knew they were acting in a simulated environment and that therefore collisions would not have been dangerous might have led to some differences compared to real driving.

In conclusion, the findings suggest that it could be beneficial for navigation systems to dynamically adapt verbal auditory messages depending on the current traffic situation. Our studies demonstrated that particular combinations of linguistic factors that determine such messages led to higher error rates in drivers’ responses than others, for example when complex messages contain redundant information. Often, however, this is only the case in specific situations, for example in busy traffic. In other situations redundant information might be helpful, for example when a message is issued well before an instruction must be executed. Consequently, the dynamic generation of navigation messages, guided by scene content, could optimise the processing of concurrent tasks and thus improve driving safety.

Acknowledgements

This research was supported by a German Science Foundation grant to the first author (DFG grant KO2350/1-1).
References


