Learning to Achieve Perfect Time Sharing: Architectural Implications of Hazeltine, Teague, & Ivry (2002)

John R. Anderson[1], Niels A. Taatgen[1], Michael D. Byrne[2]

1. Psychology Department, Carnegie Mellon University, Pittsburgh, PA 15213

2. Psychology Department, Rice University, Houston, TX 77005

Abstract

Hazeltine, Teague, & Ivry (2002) have presented data that have been interpreted as evidence against a central bottleneck and in favor of a unlimited capacity central processor as in the EPIC (Meyer & Kieras, 1997) theory. We describe simulations of their Experiment 1 and 4 in the ACT-R cognitive architecture that does possess a central bottleneck in production execution. The simulation model is capable of accounting for the emergence of near perfect time sharing in their Experiment 1 and the detailed data on the distribution of response times from their Experiment 4. With practice the central bottleneck in ACT-R will be reduced to a maximum of 50 ms (one production cycle) and can often be much less depending on timing of stages and variability in their times. We also show with a mathematical analysis of Hazeltine et al's Experiment 2, that the expected dual costs for these kinds of highly practiced tasks will be small in all circumstances, often under 10 ms.

Usually, people find it more difficult to perform two tasks at once even when the tasks involve different perceptual and response modalities. Such difficulties are often taken as evidence for a central bottleneck (Pashler, 1994; Welford, 1952). Recently, however, Schumacher et al (2001) provided evidence that with enough practice and with enough incentive participants could come to perform an aural-vocal task and a visual-manual task simultaneously with very little cost. This research was taken as evidence for the Meyer & Kieras (1997) EPIC theory that postulates central cognition is controlled by a parallel production system that is not subject to capacity limitations. More recently Hazeltine, Teague, and Ivry (2002) followed up Schumacher et al with a more extensive series of studies that addresses some possible questions about the original research and also concludes that there is very little if any central bottleneck after extensive practice. It is the research in the Hazeltine et al paper and its implications that will be the focus of this paper.

We (Byrne & Anderson, 2001) published a model showing that the basic Schumacher results could be accommodated in the ACT-R theory (Anderson & Lebiere, 1998), a production system that postulates production rule execution is serial and therefore constitutes a central bottleneck. Our purpose in this paper will be to show that the ACT-R theory is compatible with the detailed data that Hazeltine et al present and that the learning mechanisms in the theory are capable of accounting for the reduction of dual-cost effects with practice. First, we will review the task used by Schumacher et al and by Hazeltine et al and the basic ACT-R model that Byrne & Anderson proposed for this task.

Then we will elaborate on how ACT-R can account for the learning results and the detailed Hazeltine et al data.

In the original Schumacher et al version of the task, which serves as the basis for the first experiment described in Hazeltine et al., participants responded to the presentation of a circle and a tone. The circle appeared in one of three horizontal locations and participants made a spatially compatible response with their right hand, pressing index, middle, or ring finger to left, middle, or right locations. The tones lasted 150 ms and were either 220 Hz, 880 Hz, or 3520 Hz and participants responded "one", "two", or "three". In the single-task condition participants did just the visual-manual task or just the aural-vocal task. In the dual-task condition both stimuli were presented simultaneously and they were asked to do both tasks simultaneously. Over many days of practice participants come to respond virtually as fast at each task in the dual-task condition as in the single-task conditions. Thus, participants were able to perform two tasks at once with virtually no cost.

Figure 1 displays the original schedule chart for the task published in Byrne & Anderson. The line labeled cognition represents production firing where each production takes 50 ms. (an assumption shared by both ACT-R and EPIC). There is one production that converts the visual stimulus into the manual response and another production that converts the tone into the speech act. The important observation is that because it takes longer to encode the sound the two productions are offset from one another and do not interfere. (Participants did take considerably longer for the aural-vocal task that the visual-manual task). The conditions in the later experiments of Hazeltine et al eliminated this convenient offset of times by introducing a delay in the presentation of the visual condition and by increasing difficulty of the visual-manual task by either by making the visual discrimination more

difficult (Experiment 2) or by making the stimulus-response mapping incompatible (Experiments 3 and 4). Despite these changes they continued to find virtually perfect time sharing. Nonetheless, we will see that the ACT-R model does a fairly good job simulating Hazeltine et al's results.

The ACT-R Model for Hazeltine et al (2002)

In this note we will not attempt an elaborate explanation of the ACT-R theory. Rather, we refer the reader to Byrne & Anderson (2001) where the perceptual-motor details are developed and to Taatgen & Anderson (2002) where the assumptions are specified about production learning. See also Anderson et al (in press) for the most current statement of the entire theory.

The production-learning model developed for ACT-R is one that takes declarative task instructions that are interpreted to perform the task and with practice converts them into production rules for directly performing the task. Early on the task is heavy in demand on central cognition to interpret these instructions but later central cognition becomes a minimal bottleneck as in Figure 1. This accounts for both the speed up in performance of the task and the elimination of much of the dual-task cost. The key to understanding the ACT-R learning model for this task is to understand the beginning and end state of the model as it learns to perform in the dual task. The model starts out with instructions for the task represented as a set of declarative instructions that can be rendered in English as:

a. When doing a pure aural block prepare to respond to the detection of an aural stimulus with the aural task instructions.

b. When doing a pure visual block prepare to respond to the detection of a visual stimulus with the visual task instructions.
c. When doing a mixed block prepare to respond to the detection of a visual stimulus with the visual task instructions and to the detection of an aural stimulus with the aural task instructions.
d. To perform the visual task translate the visual location into a key, press that key, and check for success.
e. To perform the aural task translate the aural tone into a word, say that word, and check for success.

In addition it has committed to memory the mappings of the locations and sounds:

f. A left location translates to the index finger of the right hand
g. A middle location translates to the middle finger of the right hand
h. A right location translates to the ring finger of the right hand
i. A low tone (220 Hz) translates to saying "one"
j. A middle tone (880 Hz) translates to saying "two"
k. A high tone (3520 Hz) translates to saying "three"

As described in Anderson et al (in press), ACT-R has general interpretative procedures for converting such declarative instructions into task behavior. Part (a) of Figure 2 illustrates the sequence of productions involved in interpreting these instructions in the mixed condition when both tasks are presented. We will step through the production rules below:

A. Set Up

1. Retrieve Instruction: This retrieves instruction (c) above.
2. Retrieve Steps: This retrieves the steps involved in that instruction, in this case preparing to respond to stimuli in both modalities.
3. Prepare Visual: This sets the system to respond with instruction (d) when the visual stimulus is encoded.
4. Prepare Aural: This sets the system to respond with instruction (e) above when the aural stimulus be encoded.
5. Ready: The system notes it is finished processing the instruction and ready to respond to a stimulus.
6. Attend Visual: This requests encoding of the visual stimulus.
7. Attend Aural: This requests encoding of the aural stimulus.

B. Perform Visual-Manual Task

1. Focus on Visual: When the location is encoded it requests retrieval of instruction in (d) above.
2. Retrieve Instruction: This retrieves instruction (d) above
3. Retrieve Steps: This retrieves the steps involved in that instruction, in this case translating the location into a finger, pressing that finger, and checking the result.
4. Translate Position: A request is made to retrieve the finger corresponding to the location.
5. Retrieve Finger: One of facts f through h is retrieved.
6. Press Finger: That finger is pressed.

7. Retrieve Assessment: In response to the step of checking, first a check is made whether the result is known[1]
8. Retrieval Failure: At this starting point in the experiment nothing can be retrieved.
9. No Problems: There is no negative feedback from the experiment
10. Subgoal to Check: Set a subgoal to check the outcome
11. Return to Goal: Return this determination to the main goal.
12. Task Successful: The task has been successfully accomplished.
13. Get Ready: The system notes it is finished processing the instruction and ready to respond to a stimulus.

C. Perform Aural-Vocal Task

1. Focus on Aural: When the location is encoded it requests retrieval of instruction in (e) above.
2. Retrieve Instruction: This retrieves instruction (e) above
3. Retrieve Steps: This retrieves the steps involved in that instruction, in this case translating the tone into a word, saying that word, and checking the result.
4. Translate Tone: A request is made to retrieve the word corresponding to the tone.
5. Retrieve Word: One of facts i through k is retrieved.
6. Say Word: That word is generated.
7. Retrieve Assessment: In response to the step of checking, first a check is made whether the result is known.
8. Retrieval Failure: At this starting point in the experiment nothing can be retrieved.
9. Subgoal to Check: Set a subgoal to check the outcome
10. No Problems: There is no negative feedback from the experiment

11. Return to Goal: Return this determination to the main goal.
12. Task Successful: The task has been successfully accomplished.
13. Get Ready: The system notes it is finished processing the instruction and ready to respond to a stimulus.

When it is a single task only one of B or C will be performed and when it is a pure block the preparation in A will be simpler. However, in all cases this is a rather laborious if logical interpretation of the instructions. As Part (a) of Figure 2 illustrates, production execution at this point in time will pose a significant serial bottleneck. All of the productions for the aural task have to wait for completion of the productions from the visual task (or vice versa – there is no requirement that the visual task be performed first). In Part (a) there are perceptual encodings and motor actions but they are not part of the critical path.

Production compilation will collapse pairs of productions together. In the limit only three productions are required to do this task. All of the productions in part A can be collapsed into a single production that responds to presentation of a tone and a location with a request to encode them. The acquired production may be paraphrased:

Production A

If the goal is to perform in a mixed block
and a tone has been sounded
and a circle has appeared
THEN encode the frequency of the sound
and encode the location of the circle
and prepare to respond to an encoding of the frequency with the aural task instructions
and prepare to respond to an encoding of the location with the aural task instructions
and note that things are ready

Similarly, all the instructions in Part B above can be collapsed into a single production that responds to the appearance of the location with an appropriate key press. This requires learning that the key press will be successful. There are three productions learned for the three locations. The one for the left location can be paraphrased:

Production B

IF the location has been encoded on the left
THEN press the index finger
    and note that things are ready

Similarly, part (C) is compressed into single productions like

Production C

IF the frequency of the tone has been encoded as 220 Hz
THEN say "one"
    and note that things are ready

Part (b) of Figure 2 illustrates the situation after production compilation. We get a situation like that in Figure 1 where the two productions for the two tasks are offset and do not interfere with one another.[2]

Part (b) of Figure 2 is as far as learning can effectively proceed. Neither of these perceptual-motor productions B nor C can be collapsed with the first preparation Production A because the preparation production makes perceptual requests that require encoding from the environment before productions B or C can fire. This is one example of how the perceptual events define the limits on

collapsing productions. On the other hand the system attempts to create a combination of the last two productions into the following:

Production B & C

IF the location has been encoded on the left
and the frequency of the tone has been encoded as 220 Hz
THEN press the index finger
and say "One"
and note that things are ready

However, as the location and sound are never encoded at the same moment in the first experiment (but see our model of Hazeltine et al's Experiment 4) this production never gets to fire. This result, which is a natural outcome of the production compilation mechanism is critical to explaining one of the Hazeltine results in their first experiment. This is that participants trained on a subset of 6 of the 9 possible combinations of 3 locations and 3 tones were able to transfer to the remaining three without showing any deficit. Separate productions are always required in this experiment to handle the two modalities and such combination rules never get to be used.

There were a number of critical parameters that determine the behavior of the models for these experiments. Among these are the timings of the operations in Figure 2. We assumed each production took an average of 50 ms., the time to encode a visual location will take 50 ms, the time to encode the auditory stimulus will take 130 ms., the time to complete a finger press will be 100 ms., and the time to trigger the voice key with an utterance will be 50 ms. The production execution time is a basic parameter of ACT-R and EPIC, the visual encoding time and manual times are close to their standard values in ACT-R (85 and 100 ms). The aural and vocal times were estimated in light of the data but are within the constraints suggested by Hazeltine et al. In addition, we assume that all times had a 100% variability (the EPIC model has 67% variability) – that is, if the mean time

was T the actual times on a trial varied uniformly between T/2 and T+T/2. For instance, production times, with a mean of 50 ms., vary between 25 and 75 ms. These assumptions, especially about variability are a bit arbitrary but they serve to establish plausible benchmarks for showing that the basic results of Hazeltine et al can be predicted within the ACT-R framework, which is not that different from the EPIC framework except for the assumption of a serial bottleneck. In addition, two parameters controlled the rate of production learning: the learning rate for production utility to be .05 and the s parameter controlling noise in utilities to be .056. The first is a standard value for many models (e.g., Anderson et al, in press) but the second was estimated to fit the learning data.

Experiment 1

Hazeltine et al performed four experiments. The first involved nine participants. Seven of these participants continued to work through the remaining three experiments. We will be concerned with modeling in detail the results of the first and fourth experiment. The first experiment followed a procedure very similar to that in the first experiment of Schumacher et al (2002). On the first day participants practiced just the single tasks. They then followed with up to seven sessions on different days in which participants performed dual-task and single-task blocks. In the dual-task blocks participants experienced 6 of the 9 possible combinations of tone-location pairs. The goal was for participants to reach the point of performing as well in the dual task as in the single task. Seven of the participants reached this goal, but the data reported for this experiment are from all nine participants. After completing this phase of the experiment participants performed two more sessions during which they had to deal with the three remaining combinations of locations and tones that had been withheld as well as the other six. Our simulation of this experiment involved seven

sessions – the first just a single task, followed by 4 sessions in which dual task blocks were intermixed with single task blocks, followed by two more sessions in which the transfer stimuli were introduced. In each session the simulation experienced the same presentation sequence as the participants. We ran 20 simulated participants, which resulted in standard errors of less than 1 ms per reported mean.

The learning results from Hazeltine et al's first experiment are illustrated in Figure 3a and the results of the simulation in Figure 3b. In their experiment and in our simulation there were two kinds of single-task trials: trials that occurred in homogeneous blocks where participants were only responsible for these items and trials that were interspersed among dual-task trials. Since there was virtually no difference between these two trial types in the data nor any difference in our simulation we collapse over these. Also there were two types of dual-task trials: those that involved the original six pairings and those that involved the new ones. As they found no difference between these two types of items and our model produced none, we also collapsed over those. Thus, Figure 3 just presented performance on single-task trials and dual-task trials for the visual-motor task and aural-vocal task. Hazeltine et al report data for three periods of the experiment – the first two sessions where dual tasks were used, the last two sessions, and the still later two transfer sessions. In our simulation we collect the means of Sessions 2 & 3, 4 & 5, and 6 & 7 to correspond to these sessions.

The simulation reproduced the overall trend of reduced differences among conditions, particularly the reduced dual-task cost. The model starts out somewhat better in the single-task aural-vocal condition and somewhat worse in the single-task visual-manual condition but ends up at close to the same point as the participants. While the correspondence is not perfect, we have reproduced the

magnitude of the learning effects (both data and simulation show approximately a 100 ms improvement) and the drop out of the dual-cost with practice (in both data and simulation the average dual cost effect drops from approx 50 ms to approx 10). Also the model predicts no difference between the new and old stimulus combinations in transfer, as was observed. Given the variability among participants contributing to this data, getting such ballpark effects is all that we would expect from the learning model. The simulation of Experiment 4 will deal with data where participants are apparently more uniform and where there is a lot more detail reported. Here will we be concerned with more precise matches. From this simulation we simply conclude that ACT-R can produce the reduction in dual cost observed in the basic paradigm.

Experiment 2

The reason why the model predicts more of a dual cost effect for the aural-vocal than the visual-manual task is that the visual-manual task can complete more quickly than the aural-vocal and therefore its central bottleneck has a better chance of occurring in a position to block the aural-vocal task. The second experiment of Hazeltine et al used a discriminability manipulation (by introducing distractor circles of a different size) to slow the visual-manual task without much effect on the dual-task cost for these material. We will not provide a detailed simulation of this experiment but rather a mathematical analysis of its potential dual cost to show perhaps more transparently why ACT-R does not predict much of a dual cost even when there is not a convenient offset in encoding times for the two tasks.

In this second experiment the times were almost identical in the hard visual discrimination condition and the aural condition. While this is probably not the exact model for each task, lets assume that both the aural and the visual encoding took the 130 ms assumed for the aural encoding in Experiment 1. This means that on average the two tasks would complete at the same time and should result in maximal interference. This is the worst-case analysis of the experiment. Should one encoding complete before the other on average (and we suspect the visual encoding was still a little quicker) there would be less interference. Also the symmetry of this assumption makes the mathematical analysis simpler.

The important complication concerns the variability in the encoding times. Although the mean encoding time is 130 ms, our model assumes a uniform distribution from 65 to 195 ms for each task and these two distributions are independent of one another. For simplicity of analysis we will assume that the central bottleneck takes a constant 50 ms but the simulation that follows for Experiment 4 will allow for variability in the central bottleneck times as well.

The following is an analysis of the delay that Task 1 will cause to the central processing of Task 2, assuming that we have reached the asymptotic state in Figure 2b where each task only requires 50 ms of central processing. Task 2 will be delayed only if its encoding (Encoding 2) finishes from 0 to 50 ms after the encoding for Task 1 (Encoding 1). If Encoding 2 finishes n (< 50) ms after the Encoding 1, its central processing will be delayed by 50-n ms. There are two cases to consider:

A. Encoding 1 completes between 65 and 145 ms leaving a full 50 ms for Encoding 2 to complete. Since the distribution is uniform, the probability that Encoding 1 will complete in this interval is 80/130 = 8/13 and the probability that Encoding 2 will

complete in the following 50 ms is 5/13. The mean delay will be 25 ms since any delay between 0 and 50 ms is equally likely.

B. Encoding 1 completes between at time t between 145 and 185 ms leaving just 185-t ms for Encoding 2 to complete at a time that will result in a delay to Task 2. The probability of Encoding 1 completing in this period is 5/13, the probability of Encoding 2 completing afterwards is on average 2.5/13 and the mean expected delay is 33.3 sec on average (requires a little calculus to establish this last number).

Thus, the total expected delay is:

$$\frac{8}{13} * \frac{5}{13} * 25ms + \frac{5}{13} * \frac{2.5}{13} * 33.3ms = 8.4ms.$$

If we assume 67% variability as in the EPIC rather than the 100% variability as in ACT-R the estimate is 11.7 ms. Hazeltine et al observed a 19 ms slowing for the auditory task and a 1 ms slowing for the visual task. Thus, the observed dual cost is in the ballpark of these estimates although there is an asymmetry suggesting the visual encoding is still completing somewhat before the auditory.

Experiment 4

To get a larger dual-cost one needs more central processing than the 50 ms single production. The third experiment introduced a stimulus-response compatibility manipulation for the visual-manual task, which as we will see does require an additional rule in the ACT-R model. Still that manipulation did not have much of an effect on the dual-task cost. The fourth experiment used both

the compatibility manipulation and involved a variation in the onset of the two tasks. Since this fourth experiment is the most complex and reports the most elaborate data we will try to model it in detail. Also, this series of experiments involved seven of the participants from the original experiment who also were in the second and third experiment. By the fourth experiment they had come to display very rapid (about 250 ms for the visual-manual task and under 300 ms for the aural-vocal task) and stable responses. Thus, it is a good data set to look for detailed matches with the simulation.

These participants, after using only compatible stimulus-response pairings in the first two experiments, were asked to be also responsible for incompatible pairings in the third experiment. These incompatible mappings involved responding to the left location with the index finger (mapping unchanged), the middle location with the ring finger (mapping changed), and the right location with the middle finger (mapping changed). One test session in Experiment 4 was performed with the compatible mappings and another test session with the incompatible mappings. Since participants were responding so rapidly, we turned learning off in our simulation and assumed it was always behaving according to the terminal model in Figure 2b (thus speeding the simulation program and allowing us to collect large numbers of observations). The way we modeled the effect of compatibility was to assume that, in the incompatible condition the participants processed the stimuli as the compatible condition and then converted their response to the incompatible response. Asymptotically, this conversion took just a single production, which made the incompatible conditions 50 ms slower – close to the observed deficit.

As an added manipulation Hazeltine et al either presented the two stimuli simultaneously or offset one from another by 50 ms. Thus, there were 6 dual-task conditions defined by whether the visual task involved a compatible mapping or not, crossed with whether there was a 50 ms head start for the visual task, or the two tasks were simultaneous, or there was a 50 ms head start for the aural task. In addition there was one single-task aural condition and two single-task visual conditions. We ran 500,000 trials in each condition to reduce the error in the estimate of ACT-R's predictions to less than 0.1 ms. This is excess precision for predicting mean times but we also wanted to predict the distribution of times. To achieve the 500,000 trials we actually created a simulation of the ACT-R simulation that just reproduced the keying timing relationships and did not have the full generality of the ACT-R model since this is hardly required for modeling this experiment.

Figure 4a compares the predictions of the ACT-R model with the data for the visual-motor task and Figure 4b compares the predictions and data for the aural-vocal task. For the visual task the model and data show a strong 50 ms effect of the compatibility manipulation. The model does predict participants will be 7 ms faster on the visual task given a head start on that visual task than a head start on the aural task while there is no significant difference in the data. On the other hand, in the case of the aural task participants average a significant 14 ms longer when the tone comes first and the model predicts 12 ms. The model predicts a disadvantage for both the aural-vocal and visual-manual task when there is a head start for the tone because the head start speeds up the aural task to the point where the visual task is more likely to interrupt. The data also show a significant 13 ms slowing of the aural-vocal task in the case of the incompatible mapping while the model predicts 9 ms. Again in the model this is because the incompatible mapping slows the visual task to the point where there is more likely to be a conflict between the two tasks.

The correspondence between model and data is close if not perfect. The model predicts that the condition where the tone sounds first will result in slower responses for both the visual-manual and the aural-vocal tasks. It also predicts a larger dual-task deficit for the aural-vocal task. While the observed deficit is larger in the case of the aural-manual, as predicted, there is basically no effect in the visual-manual task, unlike the prediction.

We get the largest dual-task cost when the aural task comes first because giving it a 50 ms head start puts it into a range where its central bottleneck is more likely to compete with the central bottleneck for the visual task. This can be seen by inspecting Figure 5 that displays the mean timing of the various operations in the six dual-task conditions. Parts (a) – (c) reflect the various compatible conditions. When there is a delay between the onset of the aural and visual stimuli (part a or c of the figure), separate productions are required to initiate the aural and visual encoding. Parts (d) – (e) reflect the incompatible condition where the visual condition requires 2 productions – B1 (which basically is the same as B and produces the compatible mapping) and B2 (which converts that mapping). In part (d) note that production B1 fires but does not complete before the aural encoding is complete. In this case a composite production B2&C fires combining the aural task and the second half of the visual task. Such composite productions never got an opportunity to fire in the compatible task (parts a – c) because there was never a point where productions for both tasks were simultaneously applicable.

Looking at Figure 5 it might seem particularly straightforward what the dual-task costs would be. The aural task is delayed 20 ms in the conditions illustrated in parts (a), (d), and (e) and not at all

delayed in conditions (b), (c), and (f). The completion of the visual task is never delayed. While this analysis would rather roughly correspond to the data, it ignores the complexities produced by the variability in the times. The maximum delay in any condition can be as large as 50 ms and as small as 0. It is also quite possible when the aural task has a head start for its production to intrude on the visual task and delay that.

Hazeltine et al report a simulation to determine how long the bottleneck could be to produce the dual tasks deficits that they observed. They estimate that the bottleneck most plausibly is in the range of 20 to 30 ms. Their simulation did not allow for the possibility of variable length of stages and distribution of costs between both tasks. However, even so their estimate is only a factor of two smaller than our maximum bottleneck cost in this task, which is the 50 ms production time, and only a factor of two larger than the predicted differences between the conditions.

It is important to realize that there is relatively little bottleneck possible in our model for the task. Bottlenecks become more significant when there is more cognition as in Figure 2a. When one realizes that participants are producing only 250 ms latencies in the compatible visual task and 300 ms latencies in the incompatible visual task and the aural task, it should be apparent that there is not much time for central cognition to intervene between cognition and action. In Byrne and Anderson we describe tasks where cognition is much more substantial and where interference is much more substantial (over a second). However, we did not provide the extensive training as in the current task. In principle with enough practice any task would be converted into one where central cognition is almost totally eliminated and where there is at most 50 ms. overlap in the central component. However, this requires converting all the knowledge and contingencies into specific

production rules. While this is more than possible in the Hazeltine et al tasks, the combinatorics for complex tasks would become overwhelming.

Our argument depends critically on the distribution of times for the two tasks. Hazeltine et al provide some data to allow us to judge how closely our model corresponds to the variability in the actual data. While Hazeltine et al cannot report data on the variability in times to reach the bottleneck, they do provide data on the variability in the difference between the completion times for the two tasks. These data are reproduced in Figure 6 for the six conditions along with our simulation. Plotted there are the probabilities that the difference between the visual and and aural completion times will fall in various 25 ms bins. Given our rather blunt assumptions about stage variability we think the correspondence between the distributions of time is stunning.

Another interesting statistic reported by Hazeltine et al concerns the correlation between the completion times for the two tasks. When the aural task comes first the correlation was .03; when the tasks are simultaneous the correlation was .20, and when the visual task came first the correlation was .24. They express puzzlement at why the correlation is different in the aural-first task. Our models correlations were -.04, .20, and .17 in the three conditions. Hence we are able to reproduce this pattern. The reason for the positive (if weak) correlation in the simultaneous and visual-first conditions is because both processes wait on the firing of the first production and will share its variability. While this is still true in the aural-first condition, the interference in the bottleneck means that when one process is fast it may interrupt and delay the other process producing a negative relationship between response times.

Conclusions

In their conclusions Hazeltine et al write "These results present a serious challenge to models of dual-task performance that postulate a unitary central bottleneck" (p. 541). We do agree that they do pose a serious challenge and we think we have shown that the ACT-R theory is up to accounting for (a) the learning trends, (b) the magnitude of the dual-task interference, and (c) the distribution of response latencies. Lest we appear to claim too much credit for ACT-R we should note that our ability to account for (b) and (c) depend critically on assumptions that we have borrowed whole cloth from EPIC. Indeed, as Byrne and Anderson have argued, in some ways ACT-R is just a more uniform version of the EPIC model extending to the central module the capacity limitations EPIC assumes for the peripheral modules.

It is unlikely that the model presented here corresponds exactly to what is happening in the participants. We offer it to show that a theory with a serial bottleneck like ACT-R is compatible with the reported data and that the ACT-R production learning mechanisms can produce the emergence of near perfect time sharing with practice. As a sign that the current model is not perfect, there are effects in the data that our model does not account for. Notable in our minds are two. First, in Figure 2 our model is unable to explain the greater speed-up that occurs with practice in the auditory single-task than the visual single-task. This may reflect some perceptual learning that ACT-R does not model. Second, our model predicts that in Figure 4 there should be an effect of stimulus onset on the visual task as well as the aural task, but there is no significant effect in the case of the visual task.

References

Anderson, J. R., Bothell, D., Byrne,M.D., Douglass, S., Lebiere, C., Qin, Y. (in press) An integrated theory of Mind. *Psychological Review*

Anderson, J. R. & Lebiere, C. (1998). The atomic components of thought. Mahwah, NJ: Erlbaum.

Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. Psychological Review, 108,847-869.

Hazeltine, E., Teague, D. & Ivry, R. B. (2002). Simultaneous dual-task performance reveals parallel response selection after practice. Journal of Experimental Psychology: Human Perception & Performance 28(3), 527-545

Meyer, D. E. & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance. Part 1. Basic mechanisms Psychological Review, 104, 2-65.

Pashler (1994) Dual-task interference in simple tasks: Data and Theory. Psychological Bulletin, 116, 220-244.

Schumacher, E. H, Seymour, T. L., Glass, J. M., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (1997). Virtually perfect time sharing in dual-task performance. Paper presented at the 38th Annual Meeting of the Psychonomic Society, Philadelphia, PA.

Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. Psychological Science, 12 (2), 101-108.

Taatgen, N. A. & Anderson, J. R. (2002). Why do children learn to say "broke"? A model of learning the past tense without feedback, Cognition. *86* (2), 123-155.

Welford, A. T. (1952) The “psychological refractory period” and the timing of high speed performance – A review and a theory. British Journal of Psychology, 43, 2-19.

Acknowledgements

John R. Anderson, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213 (412) 268-2788, ja@cmu.edu. This research was supported by NASA grant NCC2-1226 and ONR grant N00014-96-01491. We would like to thank Eliot Hazeltine for his comments on an earlier draft. Correspondence concerning this article should be addressed to John R. Anderson, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213. Electronic mail may be sent to ja@cmu.edu.

Figure Captions

Figure 1 The ACT-R schedule chart for Schumacher, et al. (2001) from Byrne and Anderson (2001): VM = for Visual-Manual task,; AV = for Auditory-Vocal task; RS = Response Selection; P = Initiate perception.

Figure 2 The operations by the ACT-R model at (a) the beginning of the learning of Hazeltine et al task and (b) the end of the learning of the task. The lengths of the boxes reflect the average time for each operation. Those operations concerned with the aural-vocal task are shaded.

Figure 3 Learning to time share: (a) Experiment 1 from Hazeltine et al (2002) and (b) ACT-R simulation

Figure 4. Comparison of data and simulation for Experiment 4 from Hazeltine et al (2002): (a) Visual-manual task and (b) Aural-vocal task.

Figure 5. Timing of the ACT-R operations in the simulation of the 6 conditions in Experiment 4 of Hazeltine et al (2002). The lengths of the boxes reflect the average time for each operation. Those operations concerned with the aural-vocal task are shaded.

Figure 6. Observed and predicted distribution of intervals between the two responses in Experiment 4 of Hazeltine et al (2002). These are computed by subtracting the reaction time for the visual task from the reaction time for the aural task. The six panels reflect the 6 conditions of the experiment.

Figure 1

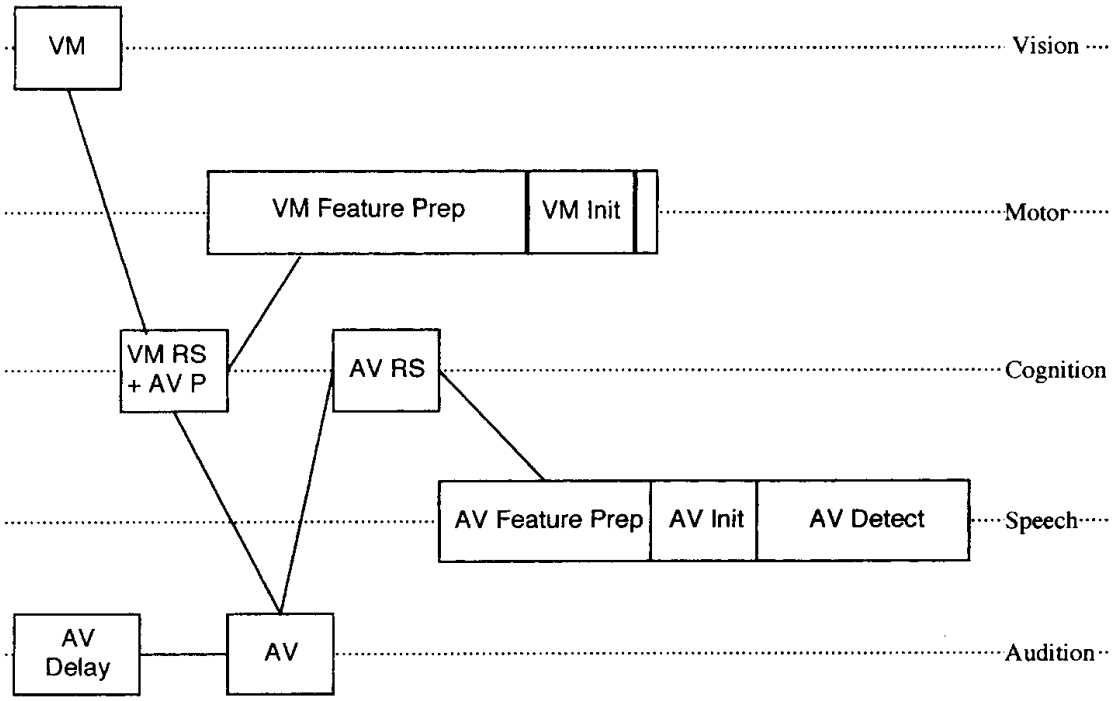

VM
Vision
VM Feature Prep
VM Init
Motor
VM RS
+ AV P
AV RS
Cognition
AV Feature Prep
AV Init
AV Detect
Speech
AV
Delay
AV
Audition

Figure 2

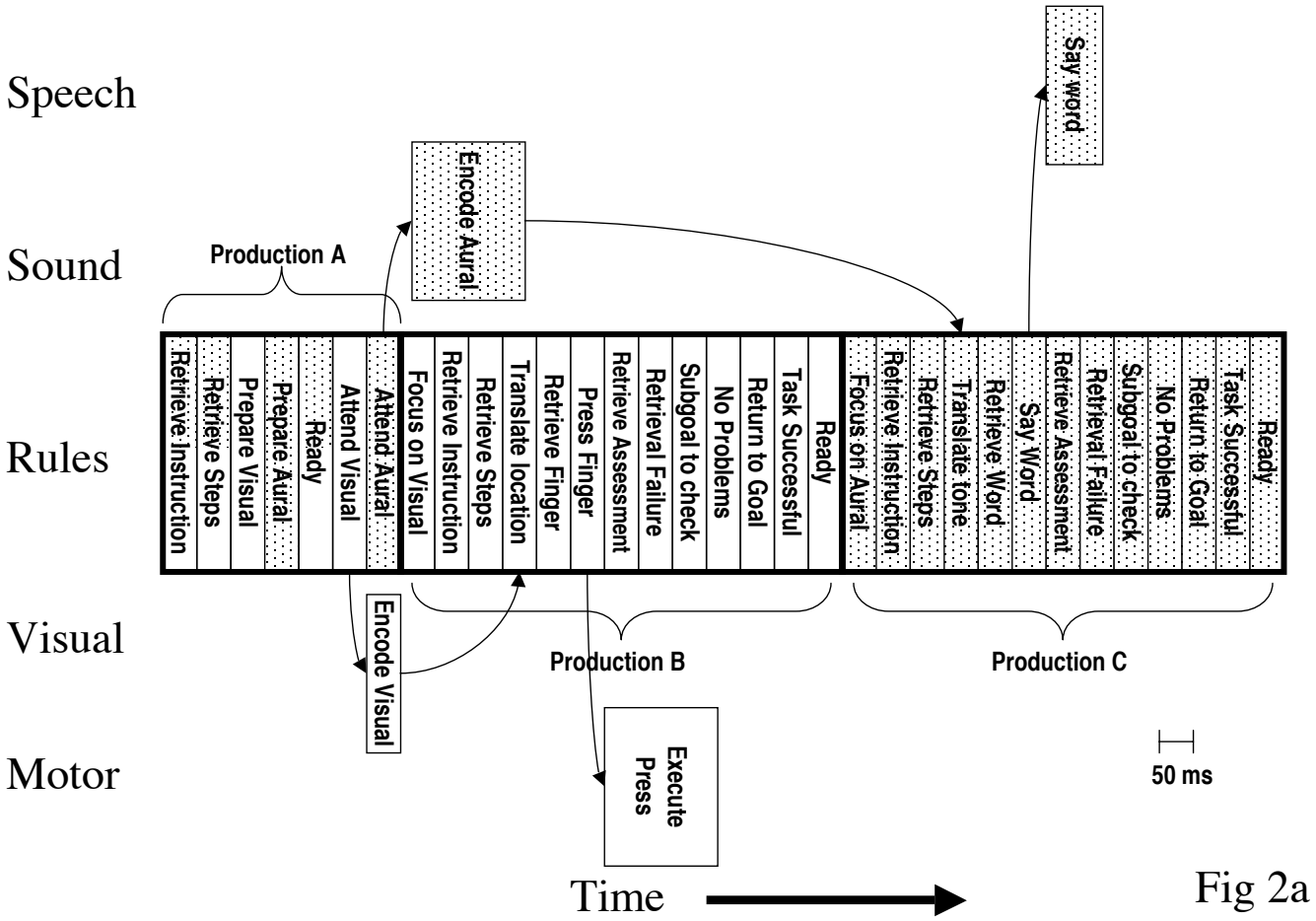


Fig 2a

Figure 2 continued

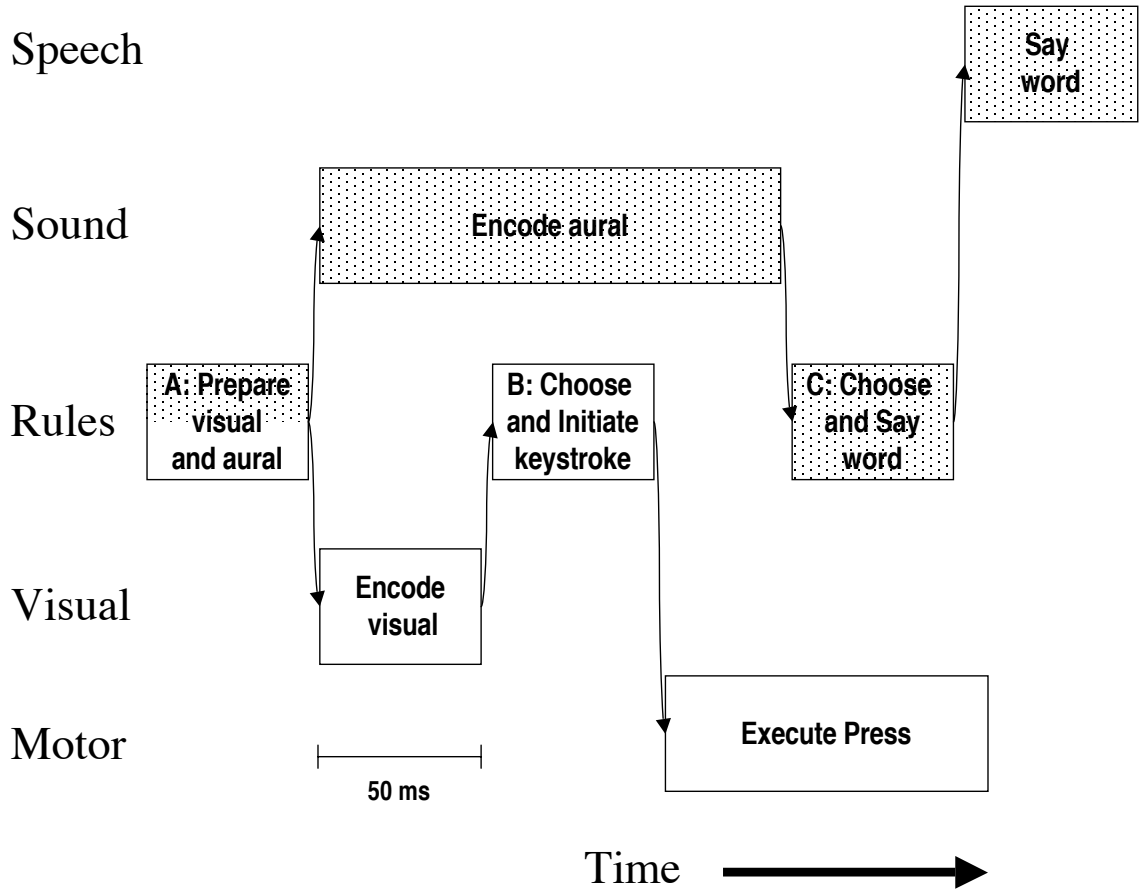


Fig 2b

Figure 3 (a) Data

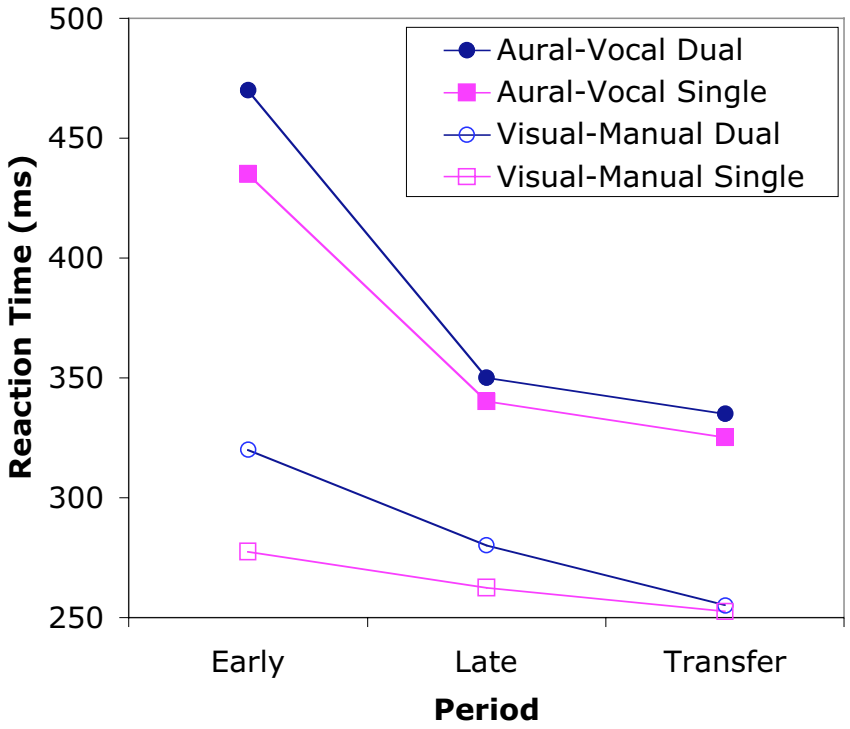


Figure 3 (b) Simulation

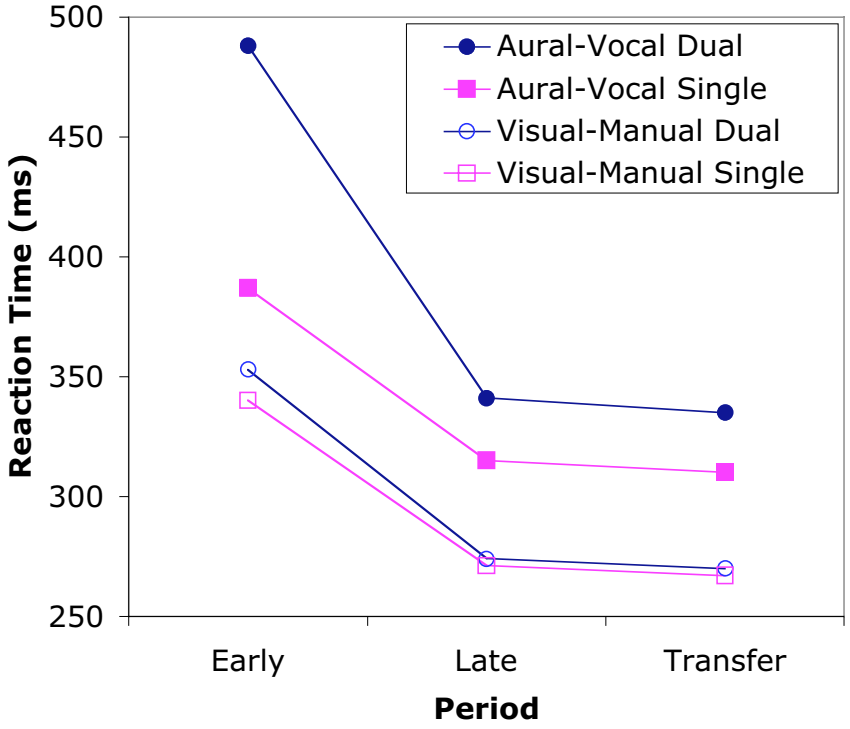


Figure 4a Visual-Manual Task

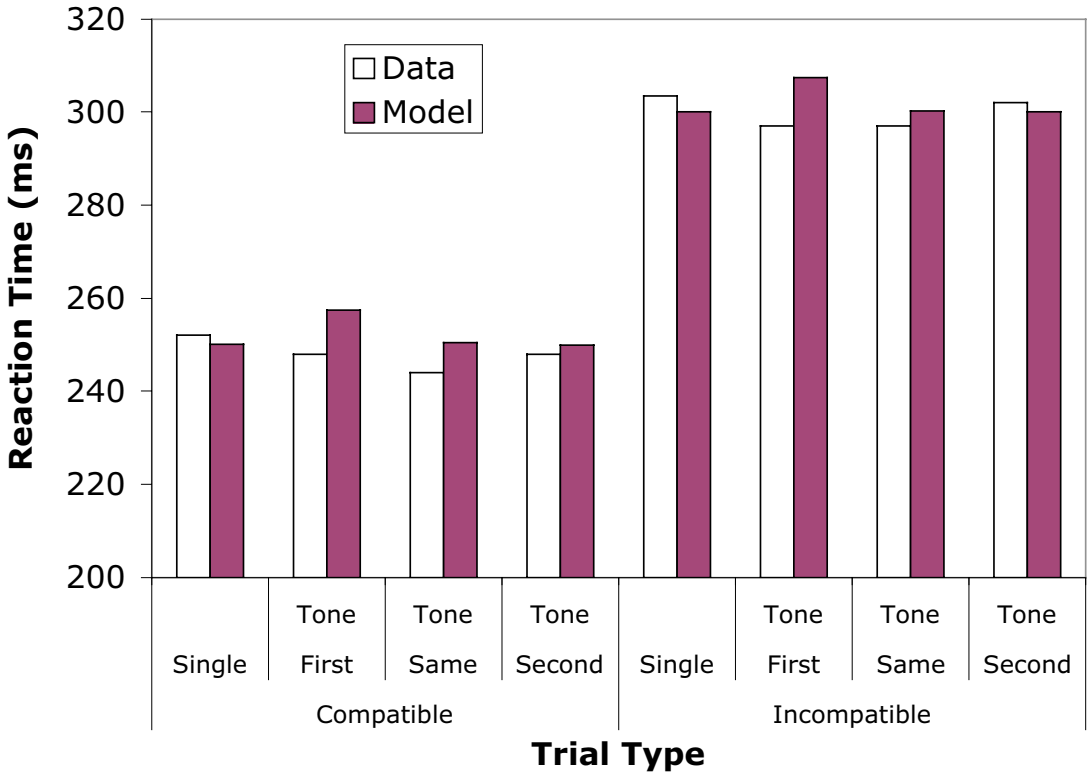


Figure 4b Aural-Vocal Task

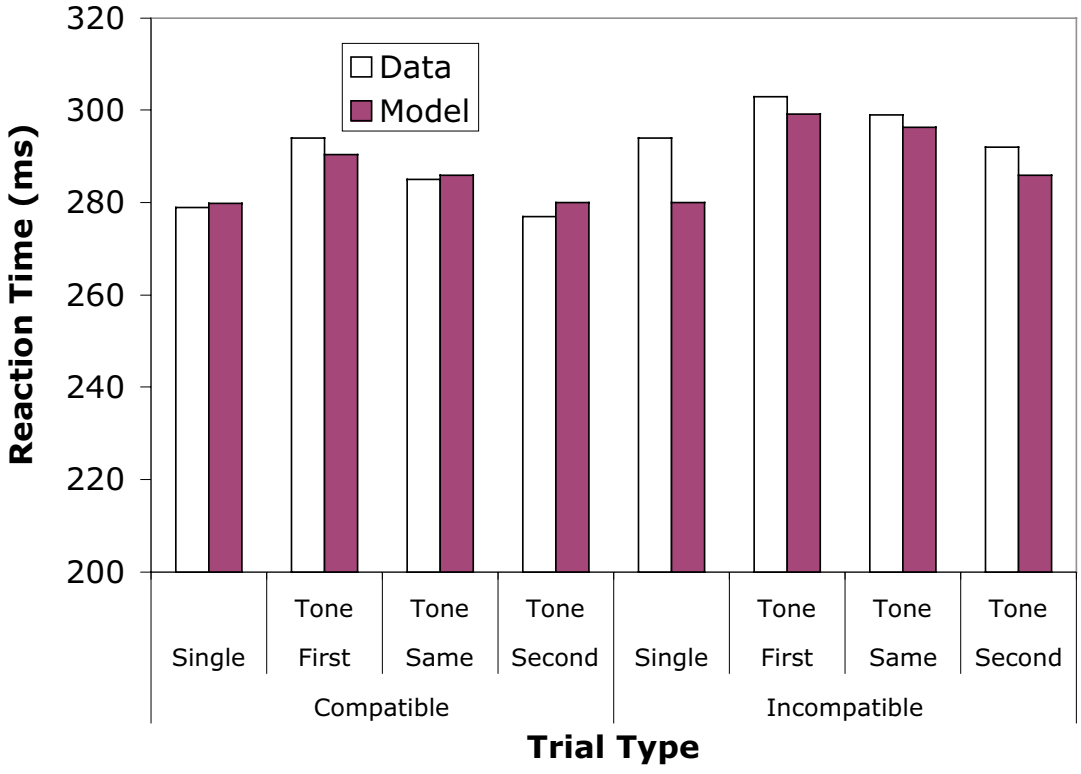


Figure 5

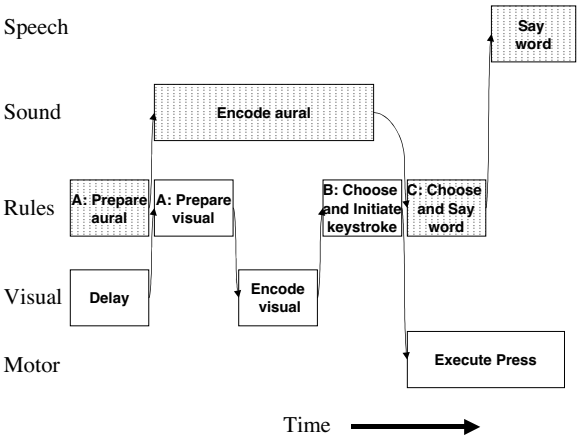


Fig 5a

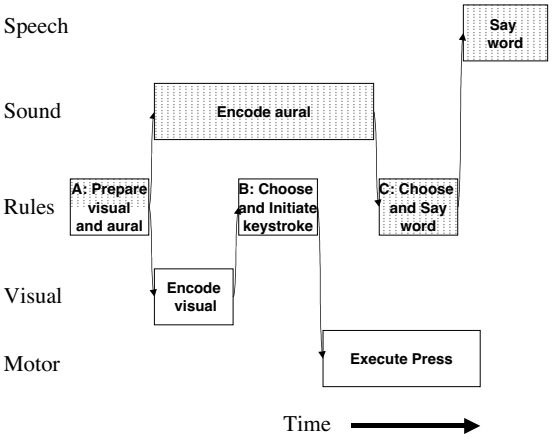


Fig 5b

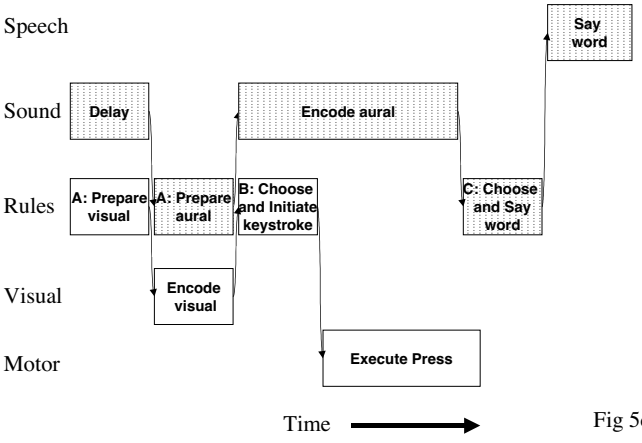


Fig 5c

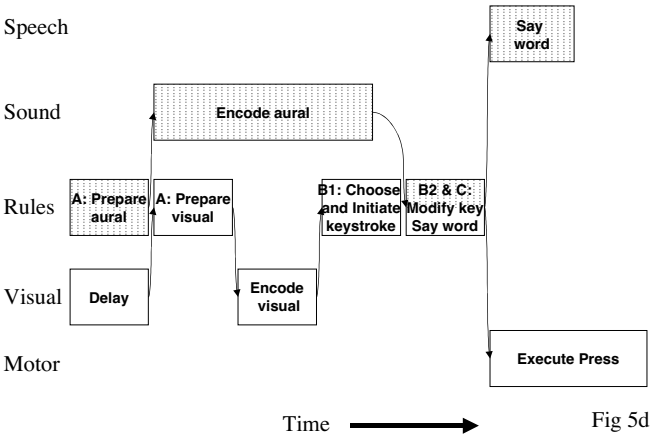


Fig 5d

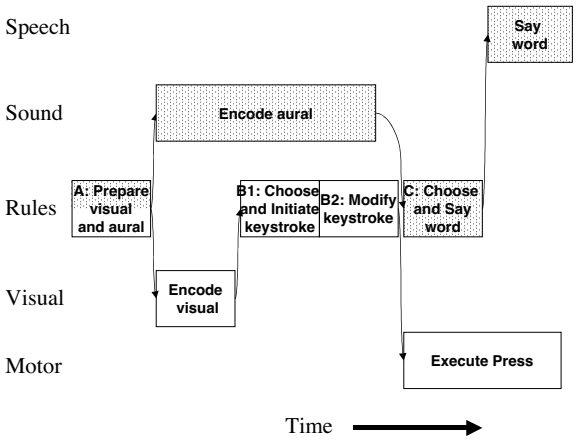


Fig 5e

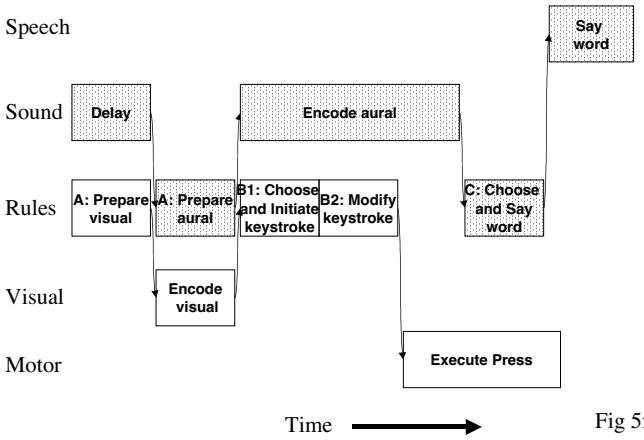
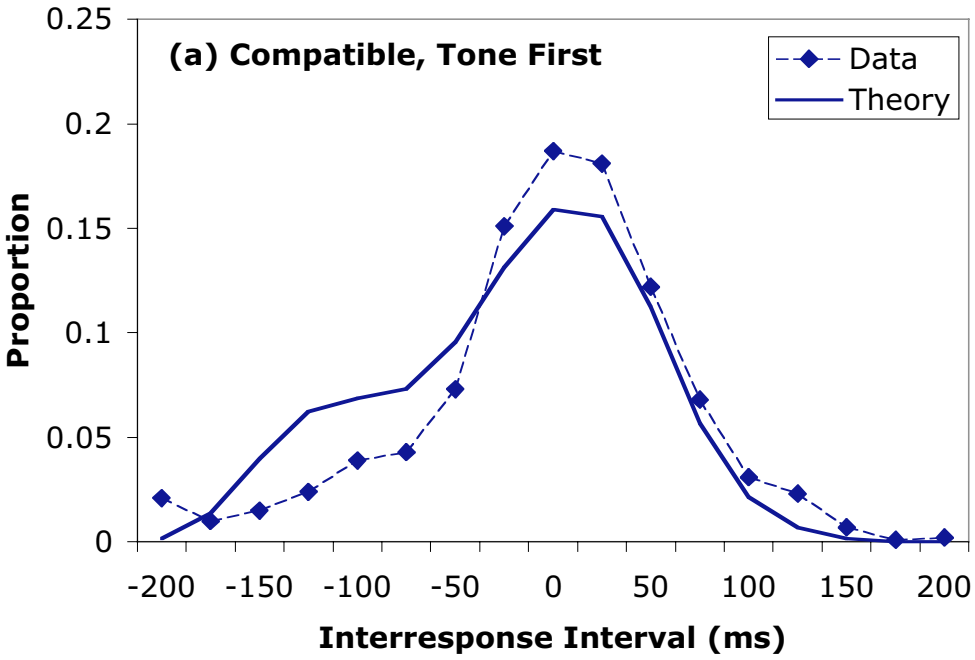


Fig 5f

Figure 6


(a) Compatible, Tone First
Data
Theory
Proportion
0.25
0.2
0.15
0.1
0.05
0
-200
-150
-100
-50
0
50
100
150
200
Interresponse Interval (ms)

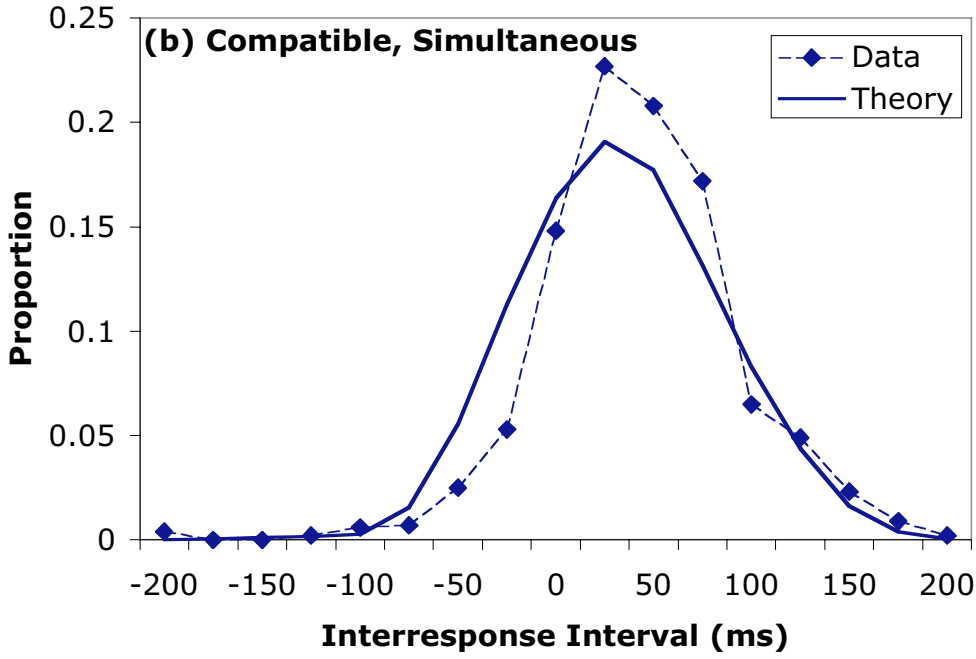
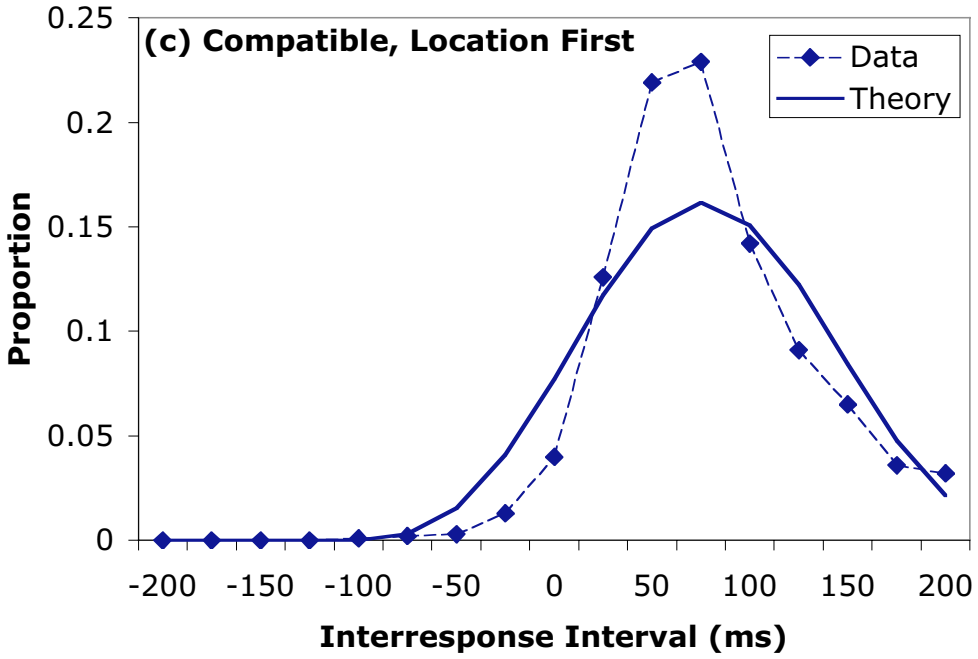

(b) Compatible, Simultaneous
Data
Theory
Proportion
0.25
0.2
0.15
0.1
0.05
0
-200
-150
-100
-50
0
50
100
150
200
Interresponse Interval (ms)

Figure 6 continued



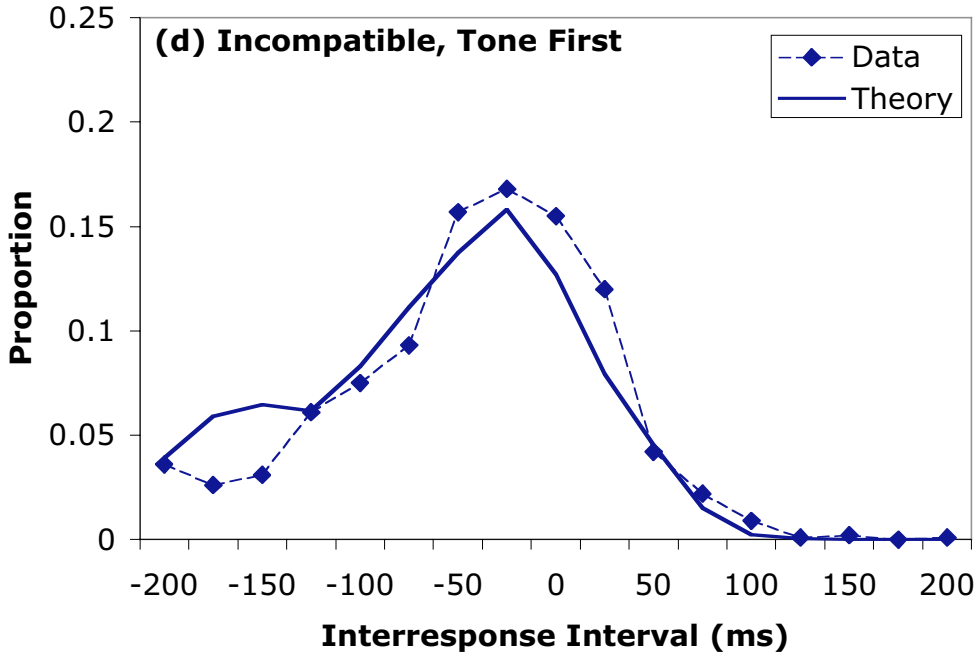
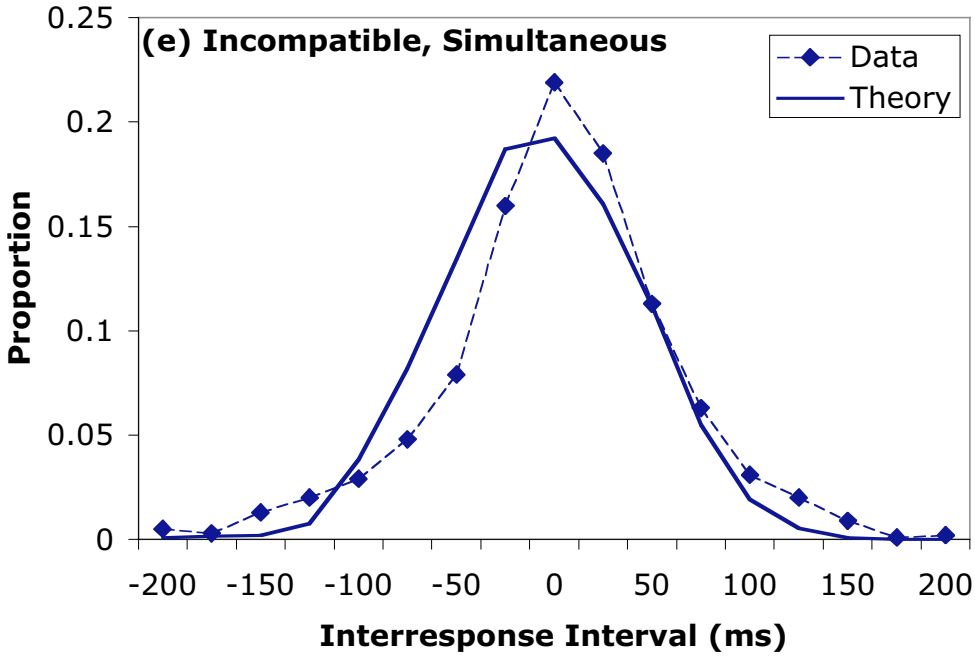
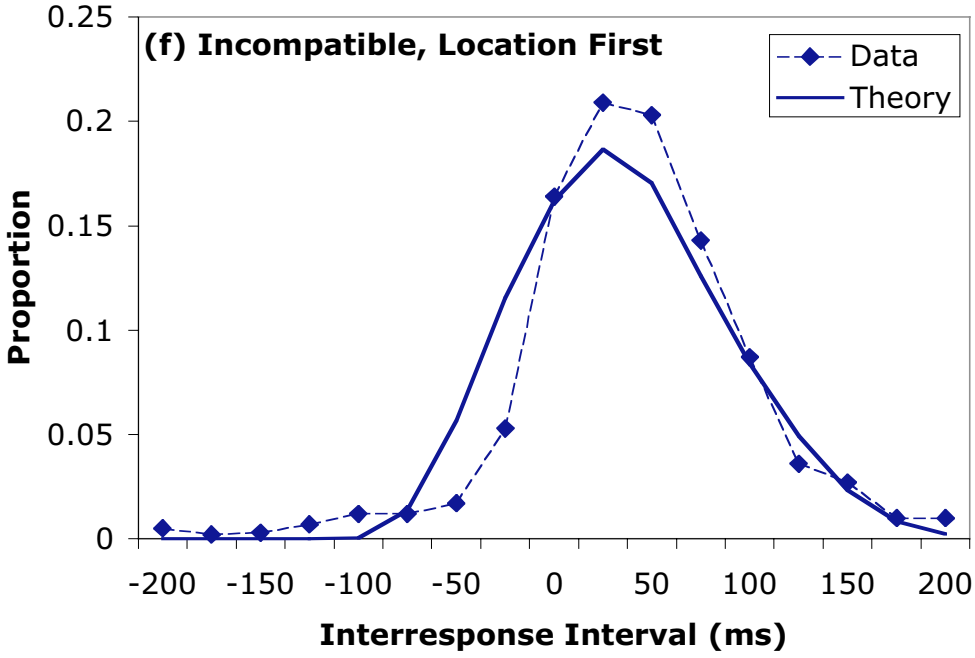


Figure 6 continued





[1] Both the visual-manual and aural-vocal task end with an assessment of whether the task has been performed successfully or not. This assessment is critical to production learning. The task can be judged successful either by setting a subgoal to judge it or by retrieving information that this action has been successful in the past. The retrieval route is tried first but initially will fail until a reliable memory is built for the outcome of such a check. This enables us to model the process by which an explicit check is dropped out and built into the learned production rules.

[2] There are some slight differences between this and Figure 1 because the production rules ACT-R learns are not identical to those Byrne and Anderson hand coded but the differences do not effect the basic explanation of perfect time sharing.