Private Speech on an Executive Task:

Relations with Task Difficulty and Task Performance

Charles Fernyhough
University of Durham

Emma Fradley
Staffordshire University

Cognitive Development MS CD282

This study was supported by two grants from Staffordshire University's SURI scheme. Some of the data were presented at the BPS Centenary Conference, Glasgow, March 2001. We thank all the children and teachers for their cheerful participation in this study; Julia Candy and Abdulrahman Al-Namlah for help with coding; Michelle Turner, Sue Gathercole, Elizabeth Meins, Sheila Ford, Holly Joseph, and two anonymous reviewers for their helpful comments. Correspondence concerning this article should be addressed to: Dr. Charles Fernyhough, Department of Psychology, University of Durham, Science Laboratories, South Road, Durham DH1 3LE, UK. Telephone +44 191 334 3243; fax +44 191 334 3241; email: c.p.fernyhough@durham.ac.uk.

Abstract

Measures of private speech and task performance were obtained for a sample of 46 5- and 6-year-olds engaged on a mechanical version of the Tower of London (ToL) task. Two different sets of 4 puzzles of increasing difficulty were attempted on 2 occasions. In line with Vygotskian predictions, there was a quadratic relation between private speech and task difficulty, but no evidence of a shift towards self-regulatory sub-types of private speech with increasing task difficulty. Levels of self-regulatory private speech were significantly related to concurrent, but not subsequent, task performance. We discuss the significance of these findings for the Vygotskian view that private speech has an adaptive function in the self-regulation of behaviour.

Keywords: private speech, self-regulation, task difficulty, task performance, Vygotsky's theory

Private Speech on an Executive Task:

Relations with Task Difficulty and Task Performance

The problem of the developmental relation between thought and language continues to present a challenge to psychologists. According to Vygotsky (1934/1987), a revolution in development is triggered when preverbal thought and preintellectual language come together to create fundamentally new forms of mental functioning. Rather than representing a developmental dead-end (Piaget, 1923/1926), the phenomenon of *private speech* (Flavell, 1966) – speech that is not obviously addressed to another listener – provides psychologists with a useful window onto this revolution in development. In Vygotsky's view, private speech represents a stage in the gradual internalisation of interpersonal linguistic exchanges whose final ontogenetic destination is inner speech, or verbal thought.

Research into the phenomenon of private speech has provided arguably the best opportunity for a thorough empirical test of Vygotsky's ideas on the relations between thought and language (Berk, 1992). Since the appearance of the first English translation of Vygotsky's work (Vygotsky, 1934/1962), psychologists have addressed several questions about the form and function of children's private speech. For example, in tracing the development of private speech over the preschool and early school years, researchers have found general support for Vygotsky's predictions about age-related changes in the incidence and structure of this form of speech. Private speech has been found to follow a predictable trajectory from overt task-irrelevant speech, to overt task-relevant speech such as self-guiding comments, and ultimately to external manifestations of inner speech such as whispering and inaudible muttering (e.g. Berk, 1986; Berk & Garvin, 1984; Bivens & Berk, 1990; Kohlberg, Yaeger, &

Hjertholm, 1968; Patrick & Abravanel, 2000; Winsler, Diaz, & Montero, 1997). There has, however, been only limited support for Vygotsky's claim that private speech 'goes underground' in the preschool and elementary school years, with some studies reporting high levels of private speech well into the elementary school years (Berk & Garvin, 1984; Berk & Potts, 1991) and into adulthood (Duncan & Cheyne, 2002).

The aim of the present study is to address two questions that can be derived from Vygotsky's views on private speech, but which have to date provided equivocal findings: the relations between private speech and task difficulty, and between private speech and task performance. We sought to address these issues by studying the private speech produced by young schoolchildren on a task with high executive demand characteristics, the Tower of London (ToL).

Private speech and task difficulty

In Vygotsky's theory, the adaptive function of private speech as a tool for self-regulation is demonstrated by the finding that, as a task becomes more difficult, children become more likely to use overt self-directed speech as an accompaniment to behaviour. Vygotsky's original (1934/1987) prediction that private speech will increase linearly with task difficulty has been supported both in studies of children (Beaudichon, 1973; Duncan & Pratt, 1997; Kohlberg et al., 1968) and young adults (Duncan & Cheyne, 2002).

The assumption of a linear relation between private speech and task difficulty has not gone unchallenged, however. In particular, it has been suggested that Vygotsky's ideas are more suggestive of a *quadratic* relation between these variables, with private speech being most likely to occur when the task is pitched within the ability range, or 'zone of proximal development' (Vygotsky, 1934/1987), of the participant (Behrend, Rosengren, & Perlmutter, 1989). When the task is too simple, self-regulatory speech will be unnecessary, because the required regulatory processes will already have been internalised. When the task is too difficult, private speech will be ineffective, resulting either in failure on the task or resort to other means of regulation.

To test this idea, Behrend et al. (1989) observed children aged between 2 and 5 engaged in solving puzzles either alone or with a parent, finding that the incidence of private speech peaked when the task was appropriate to, or slightly in advance of, the child's current level of ability. However, the non-continuous measure of difficulty used in this study, particularly the fact that the puzzles were qualitatively different from each other and were therefore not equated for perceptual complexity, meant that this quadratic relation could not be tested for statistical significance. In addition, Behrend et al.'s coding scheme for private speech, which simply distinguished between social and private speech without analysing the latter into its functional sub-types, meant that these researchers were unable to investigate how task difficulty related to the content and function of private utterances.

Private speech and task performance

Our second main focus of interest concerns the relation between private speech and task performance. The rationale behind the Vygotskian approach to this question is that, if private speech has a positive role to play in self-regulation, its use should be associated with superior task performance. Initial investigations of this issue led to equivocal results, with some researchers reporting positive associations between private speech and task performance (e.g. Beaudichon, 1973; Goodman, 1981; Winsler, Diaz, McCarthy, Atencio, & Chabay, 1999), and others reporting little or no association (e.g. Berk, 1986; Frauenglass & Diaz, 1985; see Berk, 1992, for a review). In their attempt to resolve this condundrum, Frauenglass and Diaz (1985) presented evidence that private speech is frequently more closely associated with concurrent task failure than with task success. The implication of this is that private speech, as a strategy which facilitates children's growing mastery over a task, will be more strongly correlated with future success than with concurrent success.

Support for this idea has come from reports that the incidence of private speech predicts subsequent rather than concurrent performance (Azmitia, 1992; Behrend, Rosengren, & Perlmutter, 1992; Bivens & Berk, 1990; Gaskill & Diaz, 1991). Winsler et al. (1997) showed that task-relevant speech predicted immediate post-test improvements on a selective attention task, in those instances where children were able to benefit from adult scaffolding. These authors noted that the speech–performance relation will be sensitive to the level of task difficulty at which the child is being required to operate. Thus, when a task is difficult, private speech is likely to be associated with task failure. When the task is brought within the child's ability range, in contrast, the speech–performance correlation is likely to be positive. Winsler et al.'s (1997) study was not directly able to test this hypothesis, however, as their design did not allow for a continuous measure of task difficulty. In addition, their performance measure was dichotomous (pass/fail) rather than continuous, necessarily reducing the power of their statistical analyses. Finally, they did not make the typical distinction between social and private speech, focusing exclusively on utterances made when children were working alone on the task, and automatically categorising such utterances as private.

A full answer to the question of the relation between private speech and task performance will therefore require a study design that can deliver both global and item-by-item speech–performance correlations, allowing the relation to be examined at varying levels of task difficulty. Filling this gap in the private speech literature was the second main aim of the study presented here.

Methodological issues

The present study set out to investigate Vygotskian hypotheses about private speech, task difficulty and task performance in the context of a task that allowed for continuous task difficulty and task performance measures. In planning the study, we were able to draw on a number of suggestions for methodological improvements made by private speech researchers (e.g. Diaz, 1992). Firstly, in light of evidence that non-communicative speech-to-self is stimulated by the presence of others (e.g. Goudena, 1987; Kohlberg et al., 1968), we ensured that a quasi-social context was preserved by having an experimenter in reasonably close proximity to the child during the problem-solving trials. Secondly, we made sure that our coding of private speech was sufficiently flexible to allow both global and item-by-item correlations of task performance, task difficulty, and incidence of the different sub-types of private speech. Thirdly, we asked class teachers to give ratings of individual children's talkativeness, thus providing us with independent corroboration of our laboratory measures of private and social speech. Fourthly, in line with other recent private speech studies (e.g., Winsler, Diaz, Atencio, McCarthy & Adams Chabay, 2000), we took measures of receptive verbal ability in order to control for the possibility that the functional significance of self-regulatory private speech might relate to the general cognitive and specific linguistic competences of the child.

Finally, our choice of task was intended to address a number of methodological issues relating to task difficulty, task performance and inter-individual variability in the incidence of private speech (Berk, 1992; Fuson, 1979). It is widely accepted that private speech is best studied when children are engaged in cognitively demanding tasks, such as the semantic task employed by Frauenglass and Diaz (1985), rather than lower-level cognitive tasks such as those requiring perceptual matching. Instead of using a task requiring semantic processing, we chose a task with a predominant executive component, the Tower of London (ToL) (Shallice, 1982). This task had previously been found to elicit high levels of private speech (Fernyhough, 1994), and to have good test–retest reliability (Bull, Espy, & Senn, 2004). As well as having been studied widely in other contexts (e.g. Langdon & Coltheart, 1999; Ozonoff & McEvoy, 1994; Pantelis et al., 1997), the ToL has the advantage over semantic tasks that task difficulty can be varied systematically, with, crucially, no variation in perceptual complexity. Furthermore, the task allows performance measures to be obtained at each level of difficulty, as well as globally, resulting in a rich body of performance data.

Our hypotheses were as follows: (1) that, in line with Vygotsky-inspired predictions, there would be a quadratic (rather than linear) relation between task difficulty and incidence of private speech; (2) that increasing task difficulty would lead to a predictable change in profile of sub-types of private speech, with a shift towards a greater incidence of self-regulatory utterances; (3) that the incidence of self-regulatory private speech would be positively related with concurrent task performance; (4) that the incidence of self-regulatory private speech would be positively related with future performance on the same task; and (5) that any speech–performance relations would be strongest on trials of intermediate complexity.

Method

Participants

Participants were 46 children aged 5–6 years (21 female; age M = 71.2 mths, SD = 3.98; range 64 – 78 mths) from two primary schools in an urban area of the

English Midlands. The demographic characteristics of the catchment areas of the two schools were similar, with children coming from a mixture of private and social housing. The percentage of children at both schools entitled to free school meals (indicating socioeconomic disadvantage) was at the national average. All children spoke English as their first language. One of the children was mixed race, and the remainder were White.

Procedures

Overview: Children participated in the study on two separate occasions a week apart. In each session they were asked to complete 4 trials of the Tower of London (ToL) task.

Setting: Children were seated at a table in a quiet corner of the school, with the female experimenter (E) sitting a short distance to one side. A portable video camera mounted on a tripod was placed opposite the child, giving a clear view of the task and the child's face. All sessions were videotaped and later coded for task-relevant activity, social and private speech, and task performance.

Tower of London (ToL task): In each testing session, participants attempted to solve 4 different positions of a mechanical version of the ToL task. Two identical copies of the ToL apparatus were used, each consisting of three pegs of different lengths inserted into a wooden base (20cm × 7cm × 2cm), and three coloured wooden balls (red, green and blue). The lengths of the pegs were such that one would accommodate three balls, one would accommodate two balls, and the smallest would only accommodate one ball (see Figure 1). A total of 12 different configurations is possible on the ToL. Two can be solved in a minimum of 2 moves, two can be solved in 3 moves, four can be solved in 4 moves, and four can be solved in 5 moves. The 4 trials in each session were presented in ascending order of difficulty, as follows: 2-

move, 3-move, 4-move and 5-move. A different group of 4 trials was presented in each session, making a total of 8 trials per participant. In each trial, the E presented the child with one of the copies of the apparatus in the 'standard' configuration, and then presented the other copy of the apparatus in one of the 12 target configurations. Presentation of the test trials was preceded by two warm-up trials, where only two balls (blue and green) were used. These warm-up trials were not coded.

The procedure for the test trials was as follows. The E invited the child to sit at the table, and then showed the child the ToL apparatus, saying '*Would you like to play this game?*' One copy of the apparatus (set up in the standard configuration) was placed in front of the child, and the other (set up in one of the target configurations) was placed on the table just out of reach of the child. The E then explained the rules of the game, as follows: '*You have to make this* (circling the apparatus nearest to the child) *look like this* (circling the second, target apparatus). *But there are some special rules you have to remember. You can only move one ball at a time. And you can only put the balls on the sticks, not on the table*.'

To ensure that children only moved one ball at a time, they were encouraged to place their free hand behind their back while performing the task. Participants were then told: '*Sometimes children like to talk aloud when they play this game. You can do that if you like. I bet in class you have to be quiet! While playing this game you can talk and say whatever you want to.*' This instruction was included in response to Frauenglass and Diaz's (1985) observation that elicitation of task-relevant private speech is maximized when children are given an explicit invitation to talk aloud to themselves when working on a task.

The two warm-up trials were then presented. The E began the test phase by saying: '*I'm going to make things different now. I'm going to add this red ball*.' The first target configuration was then prepared, and the E said: '*Now, can you make this* (circling the standard apparatus) *look like this?'* (circling the target apparatus). Timing for performance measures began as soon as the E had finished this last utterance. The same time-point served as the starting point for subsequent videotape coding of children's speech.

If a child became stuck, distracted or upset, the E intervened, resetting the puzzle if necessary. In such instances, which were rare (less than 2% of trials), only the second attempt at the problem was coded. At all other times the E sat quietly to one side, answering any direct questions but otherwise keeping her involvement to a minimum.

Receptive verbal ability: After the Session 2 ToL trials, participants' receptive verbal ability was assessed using the British Picture Vocabulary Scale (Dunn, Dunn, Whetton, & Pintile, 1982).

Measures

Task performance: Performance on the ToL was assessed in both testing sessions. Two measures of task performance were used: *moves to solution (MTS)*, and *time to solution (Time)*. In situations where the child ignored the rule about only moving one ball at a time, the illegitimate moves were included in the *MTS* tally. It was not deemed necessary to introduce any arbitrary penalty system to compensate for illegitimate moves; instead, it was assumed that the difficulties of children who were unable to follow the rule would be apparent from relatively higher *MTS* and *Time* scores.

ToL performance was also considered globally, collapsing across all four trials in each session. Two measures of global performance were used: (i) total time to solution (*Total Time*), and (ii) the total move-value of trials solved correctly, in the minimum number of moves, without rule-breaking (*Total Move-Value*). This latter score is identical to that used by Bull et al. (2004) for scoring performance on the ToL, and similar to that of Bishop, Aamodt-Leeper, Creswell, McGurk, & Skuse (2001). For example, a child who solved one 2-move, one 3-move and one 5-move trial in the minimum number of moves without rule-breaking would receive a score of 10 (2+3+5). No score would be given for the failed 4-move trial. The maximum possible *Total Move-Value* score was therefore 14. This score is appropriate as it indicates the number of trials correctly solved, while also giving greater credit for solution of the more complex trials.

Private speech: Private speech was coded from videotapes of Sessions 1 and 2 by an independent trained rater who was blind to the study's hypotheses. The coding process involved assigning each utterance to a superordinate category (social or private), and then further sub-dividing private utterances according to overtness and task relevance. An utterance was defined as a unit of speech containing no temporal or semantic discontinuities, where a temporal discontinuity was defined as a pause of at least 2 seconds, and a semantic discontinuity included any change of content, whether or not preceded by a pause. Utterances were then classified as social or private according to the following objective criteria (C = child, E = experimenter; Fernyhough & Russell, 1997; adapted from Diaz, 1992; Furrow, 1992; Goudena, 1992). An utterance was classified as social if: (1) *Eye contact*: C showed sustained eye contact with E during or within 2 seconds of an utterance; (2) *Behavioural*: C's behaviour involved E (through physical contact, gaze direction, etc.), or E's behaviour involved C, within 2 seconds of the utterance; (3) *Content markers*: the utterance had the same topic as E's preceding utterance, or was a question directed to E, or contained a vocative or E's name; (4) *Temporal contiguity*: the utterance occurred less than 2s after any other social utterance. Any utterance that did not meet any of these criteria for social speech was classified as private.

Each private utterance was further assigned to one of the three categories of Berk's (1986) coding scheme: *Level 1 private speech* (PS1) (task-irrelevant private speech, including word play and repetition, task-irrelevant affect expression, comments to absent, imaginary or nonhuman others); *Level 2 private speech* (PS2) (task-relevant externalised private speech, including describing one's own activity and self-guiding comments, self-answered questions, task-relevant affect expression); *Level 3 private speech* (PS3) (task-relevant external manifestations of inner speech, including inaudible muttering and whispering, and silent, verbal lip and tongue movements). A randomly selected 16% of the videotapes was coded by a naïve independent trained rater. Inter-rater reliability for the four-way distinction between social speech, PS1, PS2 and PS3 was $\kappa = 0.80$. Speech measures were divided by time spent on the task to give rate scores for each category (utterances per minute).

Teachers' ratings of talkativeness: In addition to these objective measures of children's speech, participants' class teachers were asked to rate each child for general talkativeness. Teachers were simply asked, '*On a scale from 0 to 10, how talkative would you say [name] is?*'

Results

Incidence of social and private speech

Table 1 shows the mean frequency measures for the different categories of social and private speech in Sessions 1 and 2, collapsed across all 4 trials. (The difference in *n*s is a result of one child failing to complete Session 2.) The overall rates of social and private speech in this study are comparable to those reported in other studies (e.g. Frauenglass & Diaz, 1985; Winsler et al., 2000). Paired samples t-tests showed no significant differences in speech rates between the two sessions, $t_{[44]} = -1.34 – 1.71$, n.s. As Table 1 shows, there was high inter-individual variability in the incidence of private speech, with standard deviations equal to or greater than group means (cf. Berk & Garvin, 1984). 89% of children produced at least one private utterance over both testing sessions.

Table 1 also shows intra-measure correlation coefficients for the four speech categories between Sessions 1 and 2. Rates of Level 3 private speech (PS3) were not significantly correlated between Sessions 1 and 2, suggesting that children who produced this form of speech in Session 1 were no longer doing so by Session 2. The inter-session correlation for Level 1 private speech (PS1) was not computable because of the absence of any such utterances in Session 1. Generally speaking, the very low incidence of Level 1 (overt, task-irrelevant) utterances in this study reflects the highly task-relevant nature of the private speech produced on this task, and is consistent with previous research with children of this age range (e.g., Winsler, de Leon, Wallace, Carlton, & Willson-Quayle, 2003).

Associations between teachers' ratings of children's talkativeness were investigated for children's Session 1 speech. Although all of these correlations were positive, none was significant ($rs_{[44]} = 0.13$ to 0.15) except for that with PS2, $r_{[44]} = 0.37$, $p < .05$, two-tailed. In addition, the overall rate of private speech was marginally significantly positively correlated with rate of social speech at Session 1, $r_{[44]} = 0.24$, $p < .06$, one-tailed. Neither BPVS scores nor age were related to any of the speech measures in this session.

Tower of London performance

Table 2 shows the mean moves-to-solution (*MTS*) and untransformed time-to-solution (*Time*) for each of the 4 trials in both testing sessions. One advantage of the

ToL is that it provides a measure of task difficulty which can be treated either continuously (Schall, Johnson, Lagopoulos, Jüptner, Jentzen, Thienel, Dittmann-Balçar, Bender, & Ward, 2003) or categorically (Baker, Rogers, Owen, Frith, Dolan, Frackowiak, & Robbins, 1996). In the analyses that follow, task difficulty is treated as a categorical variable in analyses of variance, with quadratic and linear models tested for significance.

A 2 (Session) × 4 (Task Difficulty) repeated measures ANOVA on the *MTS* scores showed no main effect of Session, $F_{[1,44]} = 2.75$, n.s., suggesting that there were no overall practice effects caused by repetition of the task in Session 2. As expected, there was a significant main effect of Task Difficulty, $F_{[3,132]} = 81.39$, $p < .001$, with mean *MTS* increasing linearly with Task Difficulty, $F_{[1,44]} = 237.65$, $p<.001$. There was also a marginally significant interaction between Session and Task Difficulty, $F_{[3,132]} = 2.60$, $p = 0.06$. Although the non-significance of this interaction means that it was not appropriate to investigate it further, inspection of Table 2 suggests that this may have been caused by faster solutions of the simplest tasks in Session 2, combined with slower solutions to the more complex tasks.

In order to reduce skewness in the distribution, *Total Time* scores were first log-transformed (cf. Langdon & Coltheart, 1999; Pantelis et al., 1997). A 2 (Session) × 4 (Task Difficulty) repeated measures ANOVA on the log-transformed *Time* scores showed a main effect of Session, $F_{[1,44]} = 42.76$, $p < .001$, with Session 2 positions being solved significantly faster. As expected, there was a main effect of Task Difficulty, $F_{[3,132]} = 65.37$, $p < .001$, with solution times increasing linearly with Task Difficulty, $F_{[1,44]} = 168.92$, $p<.001$. There was also a significant interaction between Session and Task Difficulty, $F_{[3,132]} = 12.75$, $p < .001$. Simple effects analyses showed the cause of this interaction to be significantly faster times to solution in Session 2 for the two simplest tasks (2-move trials, $t_{[45]} = 4.17$, $p < .001$; 3-move trials, $t_{[45]} = 7.61$, $p < .001$). This suggests that any practice effects caused by the repetition of the task in Session 2 were restricted to the simpler (2-move and 3-move) trials.

The global task performance measures, *Total Time* and *Total Move-Value*, for Session 1 showed modest associations with age ($r_{[44]} = -0.14$, n.s., and $r_{[44]} = 0.31$, $p < .05$, two-tailed, respectively) but not with BPVS ($r_{[44]} = 0.07$, n.s., and $r_{[44]} = -0.18$, n.s., respectively). Note that the (non-significant) negative association between *Total Time* and age indicates a *positive* age–performance relation, with older children recording faster times to solution. The lack of any association between receptive verbal ability and either speech measures or ToL performance meant that BPVS scores were not considered further in subsequent analyses. Correlations within the global task performance measures across Sessions 1 and 2 were computed, partialling out age. For log-transformed *Total Time*, $r_{[42]} = 0.32$, $p < .05$; for *Total Move-Value*, $r_{[42]} = 0.15$, n.s. (two-tailed tests). At Session 1, the two global performance measures were highly correlated with each other (again partialling out age), $r_{[43]} = -0.55$, $p < .001$, with longer times to solution being associated with a lower total move-value for correct solutions. There was no correlation between the two measures at Session 2, $r_{[42]} = -0.23$, n.s. (two-tailed tests).

Private speech and task difficulty

In line with Behrend et al.'s (1989) reading of Vygotsky's theory, Hypothesis 1 predicted a quadratic relation between self-regulatory private speech and task difficulty, with the highest incidence of such speech expected on tasks that were neither too easy nor too difficult for the child. It is important to note that baseline measures of performance on this task were not taken, meaning that it was impossible to determine *a priori* which tasks would be easy or difficult for any particular child.

That said, one advantage of the ToL is that judgements about difficulty can be made on the basis of analyses of the cognitive demands of individual task trials (e.g., Bull et al., 2004) rather than on post hoc assessments of performance. In light of previous research with this task and this age-group (e.g. Bull et al., 2004), we had reason to assume that the ToL trials would fall within the general competence range of our sample. In support of this assumption, performance on the simplest trials was generally very high, and none of the children in our study appeared unable to get started on the more difficult trials.

In order to investigate changes in speech rates with increasing task difficulty, a 2 (Session) × 4 (Task Difficulty) × 4 (Speech Type) ANOVA was conducted with speech rate as the dependent variable. There was no main effect of Session, $F_{[1,44]} < 1$, n.s. There were main effects of Task Difficulty, $F_{[3,132]} = 4.56$, $p < .01$, and Speech Type, $F_{[3,132]} = 12.30$, $p < .001$, and no interactions. The effect of Task Difficulty was found to fit a quadratic model, $F_{[1,44]} = 8.73$, $p < .005$ (for the linear model, $F < 1$, n.s.). These results thus provide support for Hypothesis 1, in that speech peaked when tasks were of intermediate levels of complexity. The lack of any significant interaction between Speech Type and Task Difficulty suggests that, contrary to Hypothesis 2, there were no changes in private speech sub-type profiles as Task Difficulty increased. Rather, the incidence of *all* forms of speech peaked when tasks were of moderate difficulty. These findings are represented graphically in Figure 2. Note that, because of the lack of any main effect of Session, speech data in this figure are collapsed across both sessions.

Private speech and task performance

1. Global correlations: To test Hypotheses 3 and 4, global correlations were computed for Session 1 and Session 2 private speech (PS2 and PS3) and task

performance measures, partialling out age. To control for differences in overall verbosity, a second covariate was partialled out, namely the rate of social speech (utterances per minute) averaged across both testing sessions. The rationale for this relates to the finding, reported above, that all speech forms peaked on moderately difficult tasks. We therefore wished to determine whether it was self-regulatory private speech specifically that was related to task performance, or use of any kind of speech, including social. The resulting correlations are shown in Table 3. Session 1 PS2 was associated with significantly faster times to solution in Session 1, but showed no association with performance in Session 2. Session 1 PS3 was associated with higher *Total Move-Value* scores in Session 1, but showed no association with performance in the subsequent session. To control for the possibility that any relation with subsequent performance might be masked by individual differences in ability on the task, these correlations were recomputed partialling out Session 1 time-to-solution. The correlations between Session 1 PS and Session 2 performance remained non-significant (*r*s = -0.04 to 0.08, n.s.). Thus, our findings show support for Hypothesis 3 (concerning concurrent speech–performance relations) but not Hypothesis 4 (namely, that use of private speech would relate to future task performance).

2. Position-by-position analyses: If assumptions about the self-regulatory role of private speech are correct, associations between private speech and task performance should be strongest on tasks of intermediate complexity (Hypothesis 5). On this reasoning, when a task is simple it can be solved without the need for self-regulatory speech, while on the most complex tasks self-regulation is unlikely to be effective. To accommodate the possibility that global correlations between private speech and task performance may mask interesting interactions with task difficulty

(e.g. Diaz, 1992, Winsler et al., 1997), separate correlation coefficients between private speech (PS2 and PS3) and time to solution were computed for each of the four difficulty levels at Session 1, partialling out age and rate of social speech (see Table 4). None of these correlations attained statistical significance. In particular, Hypothesis 5 would predict that speech–performance correlations on the simplest (2-move) and most complex (5-move) tasks would be close to zero, while those for the intermediate tasks (3-move and 4-move) would be significantly negative (indicating positive relations between speech and performance). No evidence for this pattern is found.

3. Item-by-item speech–performance combinations: Winsler et al. (2000) argue that each task-item in private speech research should be analysed separately for the coincidence of task-relevant speech with task success or failure. Following this example, each trial in each session was coded according to the type of speech that accompanied it, and whether it resulted in success or failure. For these purposes, task success was defined as solving the puzzle in the minimum number of moves without rule-breaking; any other outcome was classified as failure. Three categories of speech accompaniment were used: *relevant* (task-relevant private speech, i.e. PS2 and PS3), *irrelevant* (task-irrelevant private speech [PS1] and social speech), and *silence*. The category *relevant* was superordinate to *irrelevant*, meaning that any incidence of task-relevant speech on an item led to a coding of *relevant*.

Table 5 shows how the frequencies of each of the six possible speech–outcome combinations varied with task difficulty. Our findings show some support for the Vygotskian view of the relation between private speech, task difficulty and task performance. For the simplest puzzle (2-move), the most frequent combination was *silence/success*, supporting the view that children are able to solve these puzzles without the use of overt private speech. For the next two levels of difficulty (3-move and 4-move), the most frequent combination was *relevant/success*. For the most complex trials (5-move), the most frequent combination was *relevant/failure*. As expected, the overall percentage of successful items decreased with task difficulty, from 86.9% for the simplest puzzles to 52.2% for the most challenging. However, because of confounding of between-subject and within-subject variables, further statistical analysis of these data is unwarranted.

Discussion

The aim of the present study was to investigate a number of hypotheses about private speech, task difficulty and task performance. By choosing a task that combined manageable outcome measures with a continuously variable level of difficulty, we were able to make a number of methodological improvements on previous private speech research. Levels of private speech were comparable to previous studies, intra-measure correlations showed the speech measures to be reasonably stable between the testing sessions, and there was evidence for an association between private and social speech levels. Class teachers' ratings of children's general talkativeness showed a modest degree of correlation with these experimental speech measures. Task performance measures were moderately well correlated with each other, and showed a predictable relation with task difficulty. Finally, the fact that the majority of children in our study engaged in at least some private speech is in line with recent findings of support for Vygotsky's conjecture that private speech represents a universal stage in development (Matuga, 2003; Winsler et al., 2003).

Our first hypothesis concerned the relation between private speech and task difficulty. Our use of a continuously variable measure of task difficulty meant that we were able to conduct a strict test of Vygotsky-inspired ideas about the relation between private speech and task difficulty, namely that the incidence of private speech will peak when the task is neither too simple nor too difficult. In line with the findings of Behrend et al. (1989), we found a quadratic relation between task difficulty and incidence of the different speech sub-types, with the highest levels of speech occurring on tasks of medium difficulty. The pattern of means was shown to fit a quadratic model, thus providing support for our first hypothesis.

Our second hypothesis was that increasing task difficulty would lead to a predictable change in private speech sub-type profiles, with a shift towards greater incidence of self-regulatory utterances (PS2 and PS3). Analyses of the private speech data across all four levels of difficulty gave no support to this prediction, with no interaction between Speech Type and Task Difficulty. Indeed, all forms of speech (including social) showed the relation with task difficulty. The lack of any interaction here may have been due in part to the very low incidence of the least developmentally advanced form of private speech, overt task-irrelevant private speech (PS1). It seems plausible that the school context in which the task was presented, and the demanding nature of the task, ensured that children's attention was task-focused from the very start. A less formal context of problem-solving might be expected to produce higher levels of task-irrelevant speech, particularly on the simpler tasks, and thus a greater likelihood of a changing profile of private speech sub-types with increasing task difficulty.

Hypotheses 3 and 4 concerned the relation between private speech and task performance. The preceding analyses showed that *all* speech forms showed the quadratic relation with task difficulty; therefore, it was essential to establish whether it was private speech *per se*, or speech in general, that was associated with task performance. First, we considered whether the incidence of PS2 and PS3 correlated with global measures of concurrent task performance, when controls were made for overall verbosity. In support of Hypothesis 3, Session 1 PS2 was significantly correlated with concurrent time-to-solution, and Session 2 PS3 was associated with total move-value scores. This provides support for the idea that it is specifically private speech, rather than social speech, that is associated with concurrent task performance. Contrary to Hypothesis 4, there was no relation between Session 1 private speech and global measures of performance in Session 2, and this lack of association was maintained even when Session 1 performance was included as a covariate. We therefore found no support for the proposal of Frauenglass and Diaz (1985) that private speech will be more positively associated with future than with concurrent performance.

Our final hypothesis was that patterns of speech–performance relations would relate to task difficulty, with the strongest associations between speech and performance appearing on tasks of intermediate complexity. Separate speech–performance correlations for each level of difficulty showed no support for this hypothesis. Although violation of statistical assumptions meant that their patterns could not be tested for statistical significance, relative frequencies of the six possible speech–outcome combinations formed a predictable pattern, with the simplest tasks most frequently associated with silence and success, and the most complex tasks tending to be associated with task-relevant private speech and failure.

The present study has thus provided useful data on Vygotskian hypotheses about the relations between private speech, task difficulty and task performance. In the main, our findings are consistent with Vygotsky's view that private speech represents a developmental waystation between social speech and inner speech. Levels of self-regulatory private speech were positively correlated with concurrent task performance, showed a quadratic relation with task difficulty, and demonstrated meaningful patterns of speech–performance relations on item-by-item analyses. One hypothesis that was not supported was the suggestion that private speech would relate more strongly to more strongly to future than to concurrent task performance. Our findings are thus at odds with the small number of studies that have reported such relations.

A closer look at these earlier studies suggests some possible reasons for this discrepancy. Firstly, Behrend et al. (1992) used undifferentiated measures of private speech, failing to distinguish between task-irrelevant and task-relevant utterances. It seems plausible that, if care is taken to consider the *content* of private utterances, relations with concurrent performance should emerge. Indeed, in a study that distinguished task-relevant from task-irrelevant speech, Azmitia (1992) found evidence that both concurrent and future performance were related to task-relevant private speech. Another study reporting relations with future but not concurrent performance (Bivens & Berk, 1990) looked at children's mathematics achievement a full year after private speech measures were taken, a much longer timescale than that proposed by Diaz (1992). Finally, Winsler et al. (1997) only considered immediate post-test improvements that occurred as a consequence of adult scaffolding, and did not make the typical distinction between social and private speech.

We suggest, then, that if care is taken to distinguish sub-types of private speech, to distinguish social from private speech, and to choose a task which is both appropriate to the elicitation of private speech and which leads to rich outcome measures, relations between private speech and task performance may become more visible. A further *desideratum* for future research is that it employ a longitudinal methodology. Such an approach has already begun to prove fruitful in this area (e.g. Winsler et al., 2000). One of the most valuable contributions that might emerge from such studies is a better understanding of the minority of children who, at any given time, appear to use no private speech at all. This may be because such children have already fully internalised private speech, or alternatively because they have not yet got started on the internalisation process described by Vygotsky.

Another challenge for future research is to determine whether, even after internalisation, overt self-regulatory private speech can have a facilitatory effect on performance. As noted in the Introduction, Vygotsky's claim that private speech 'goes underground' during the early school years has received only modest support (e.g. Berk & Garvin, 1984; Berk & Potts, 1991). One possibility is that, even after children have largely internalised private speech, they continue to employ overt self-regulation when the task is sufficiently demanding. Such re-emergence (Fernyhough, in press) of already-internalised private speech might help to explain the apparent discrepancy between the very early appearance of private speech (Furrow, 1984; Furrow, 1992), and its persistence into later childhood and adulthood. Rather than providing a window onto internalisation, it is possible that, for some individuals, private speech in problem-solving contexts reflects a movement from internal to external self-regulation which can temporarily reverse transitions made earlier in development. Again, only careful longitudinal research can answer this question satisfactorily.

References

Azmitia, M. (1992). Expertise, private speech, and the development of self-regulation. In R. M. Diaz & L. E. Berk (Eds.), *Private speech: From social interaction to self-regulation* (pp. 101-122). Hove, UK: Lawrence Erlbaum Associates.

Baker, S. C., Rogers, R. D., Owen, A. M., Frith, C. D., Dolan, R. J., Frackowiak, R. S. J., & Robbins, T. W. (1996). Neural systems engaged by planning: A PET study of the Tower of London task. *Neuropsychologia, 34*, 515-526.

Beaudichon, J. (1973). Nature and instrumental function of private speech in problem solving situations. *Merrill-Palmer Quarterly, 19*, 119-135.

Behrend, D. A., Rosengren, K. S., & Perlmutter, M. (1989). A new look at children's private speech: The effects of age, task difficulty, and parent presence. *International Journal of Behavioral Development, 12*, 305-320.

Behrend, D. A., Rosengren, K. S., & Perlmutter, M. (1992). The relation between private speech and parental interactive style. In R. M. Diaz & L. E. Berk (Eds.), *Private speech: From social interaction to self-regulation* (pp. 85-100). Hove, UK: Lawrence Erlbaum Associates.

Berk, L. E. (1986). Relationship of elementary school children's private speech to behavioral accompaniment to task, attention, and task performance. *Developmental Psychology, 22*, 671-680.

Berk, L. E. (1992). Children's private speech: An overview of theory and the status of research. In R. M. Diaz & L. E. Berk (Eds.), *Private speech: From social interaction to self-regulation* (pp. 17-53). Hove, UK: Lawrence Erlbaum Associates.

Berk, L. E., & Garvin, R. A. (1984). Development of private speech among low-income Appalachian children. *Developmental Psychology, 20*, 271-286.

Berk, L. E., & Potts, M. K. (1991). Development and functional significance of private speech among attention-deficit hyperactivity disordered and normal boys. *Journal of Abnormal Child Psychology, 19*, 357-377.

Bishop, D. V. M., Aamodt-Leeper, G., Creswell, C., McGurk, R., & Skuse, D. H. (2001). Individual differences in cognitive planning on the Tower of Hanoi task: Neuropsychological maturity or measurement error? *Journal of Child Psychology and Psychiatry, 42*, 551-556.

Bivens, J. A., & Berk, L. E. (1990). A longitudinal study of the development of elementary school children's private speech. *Merrill-Palmer Quarterly, 36*, 443-463.

Bull, R., Espy, K. A., & Senn, T. A. (2004). A comparison of performance on the Towers of London and Hanoi in young children. *Journal of Child Psychology and Psychiatry, 45*, 743-754.

Diaz, R. M. (1992). Methodological concerns in the study of private speech. In R. M. Diaz & L. E. Berk (Eds.), *Private speech: From social interaction to self-regulation* (pp. 55-81). Hove, UK: Lawrence Erlbaum Associates.

Duncan, R. M., & Cheyne, J. A. (2002). Private speech in young adults: Task difficulty, self-regulation, and psychological predication. *Cognitive Development, 16*, 889-906.

Duncan, R., & Pratt, M. (1997). Microgenetic change in the quality and quantity of preschoolers' private speech. *International Journal of Behavioral Development, 20*, 367-383.

Dunn, L. M., Dunn, L. M., Whetton, C., & Pintilie, D. (1982). *British Picture Vocabulary Scale*. Windsor, UK: NFER-Nelson.

Fernyhough, C. (1994). *Social and private speech as determinants of early cognitive functioning*. Unpublished PhD thesis, University of Cambridge.

Fernyhough, C. (in press). Alien voices and inner dialogue: Towards a developmental account of auditory verbal hallucinations. In press, *New Ideas in Psychology*.

Fernyhough, C., & Russell, J. (1997). Distinguishing one's own voice from those of others: A function for private speech? *International Journal of Behavioral Development, 20*, 651-665.

Flavell, J. H. (1966). Le langage privé. *Bulletin de Psychologie, 19*, 698-701.

Frauenglass, M. H., & Diaz, R. M. (1985). Self-regulatory functions of children's private speech: A critical analysis of recent challenges to Vygotsky's theory. *Developmental Psychology, 21*, 357-364.

Furrow, D. (1984). Social and private speech at two years. *Child Development, 55*, 355-362.

Furrow, D. (1992). Developmental trends in the differentiation of social and private speech. In R. M. Diaz & L. E. Berk (Eds.), *Private speech: From social interaction to self-regulation* (pp. 143-158). Hove, UK: Lawrence Erlbaum Associates.

Fuson, K. C. (1979). The development of self-regulating aspects of speech. In G. Zivin, (Ed.), *The development of self-regulation through private speech* (pp. 135-217). New York: Wiley.

Gaskill, M. N., & Diaz, R. M. (1991). The relation between private speech and cognitive performance. *Infancia y Aprendizaje, 53*, 45-58.

Goodman, S. (1981). The integration of verbal and motor behavior in preschool children. *Child Development, 52*, 280-289.

Goudena, P. P. (1987). The social nature of private speech of preschoolers during problem solving. *International Journal of Behavioral Development, 10*, 187-206.

Goudena, P. P. (1992). The problem of abbreviation and internalization of private speech. In R. M. Diaz & L. E. Berk (Eds.), *Private speech: From social interaction to self-regulation* (pp. 215-224). Hove, UK: Lawrence Erlbaum Associates.

Kohlberg, L., Yaeger, J., & Hjertholm, E. (1968). Private speech: Four studies and a review of theories. *Child Development, 39*, 691-736.

Langdon, R., & Coltheart, M. (1999). Mentalising, schizotypy, and schizophrenia. *Cognition, 71*, 43-71.

Matuga, J. M. (2003). Children's private speech during algorithmic and heuristic drawing tasks. *Contemporary Educational Psychology, 28*, 552-572.

Ozonoff, S., & McEvoy, R. (1994). A longitudinal study of executive function and theory of mind development in autism. *Development and Psychopathology, 6*, 415-431.

Pantelis, C., Barnes, T. R. E., Nelson, H. E., Tanner, S., Weatherley, L., Owen, A. M., & Robbins, T. W. (1997). Frontal-striatal cognitive deficits in patients with chronic schizophrenia. *Brain, 120*, 1823-1843.

Patrick, E., & Abravanel, E. (2000). The self-regulatory nature of preschool children's private speech in a naturalistic setting. *Applied Psycholinguistics, 21*, 45-61.

Piaget, J. (1926). *The language and thought of the child*. London: Routledge and Kegan Paul. Original work published 1923.

Schall, U., Johnston, P., Lagopoulos, J., Jüptner, M., Jentzen, W., Thienel, R., Dittmann-Balçar, A., Bender, S., & Ward, P. B. (2003). Functional brain maps of Tower of London performance: A positron emission tomography and functional magnetic resonance imaging study. *Neuroimage, 20*, 1154-1161.

Shallice, T. (1982). Specific impairments in planning. *Philosophical Transactions of the Royal Society of London, B298*, 199-209.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press. Original work published 1934.

Vygotsky, L. S. (1987). *Thinking and speech*. In *The Collected Works of L. S. Vygotsky, Vol. 1*. New York: Plenum. Original work published 1934.

Winsler, A., Diaz, R. M., McCarthy, E. M., Atencio, D. J., & Chabay, L. A. (1999). Mother-child interaction, private speech, and task performance in preschool children with behavior problems. *Journal of Child Psychology and Psychiatry, 40*, 891-904.

Winsler, A., Diaz, R. M., Atencio, D. J., McCarthy, E. M., & Chabay, L. A. (2000). Verbal self-regulation over time in preschool children at risk for attention and behavior problems. *Journal of Child Psychology and Psychiatry, 41*, 875-886.

Winsler, A., Diaz, R. M., & Montero, I. (1997). The role of private speech in the transition from collaborative to independent task performance in young children. *Early Childhood Research Quarterly, 12*, 59-79.

Winsler, A., de Leon, J. R., Wallace, B. A., Carlton, M. P., & Willson-Quayle, A. (2003). Private speech in preschool children: Developmental stability and change, across-task consistency, and relations with classroom behaviour. *Journal of Child Language, 30*, 583-608.

Table 1: Mean Frequency Measures of Social and Private Speech at Sessions 1 and 2 (utterances per minute)

	Session 1 (*n*=46)	Session 2 (*n*=45)	*r*s
Social speech	0.65 (0.90)	0.41 (0.69)	0.27*
Private speech	2.43 (2.42)	2.54 (3.10)	0.44**
PS1 (overt, task-irrelevant)	0.0 (0.0)	0.03 (0.14)	-
PS2 (overt, task-relevant)	1.38 (2.00)	1.47 (2.61)	0.55***
PS3 (covert, task-relevant)	1.05 (1.21)	1.04 (1.33)	0.10

Figures in brackets are *SD*. Figures in the third column are intra-measure correlations between Sessions 1 and 2, d.f.= 43, two-tailed tests. *p<.07, **p<.005, ***p<.001

Table 2: Mean Moves-to-Solution (MTS) and Untransformed Time-to-Solution (sec) for Each of the Four Trials at Both Testing Sessions

	Level of Difficulty			
	2-move	3-move	4-move	5-move
MTS (Session 1)	3.00 (4.26)	3.26 (0.65)	5.22 (3.8)	7.26 (3.26)
MTS (Session 2)	2.04 (0.29)	3.61 (3.40)	5.76 (3.59)	9.09 (5.44)
Time (Session 1)	11.07 (12.4)	46.35 (47.8)	33.41 (38.1)	42.07 (51.9)
Time (Session 2)	5.91 (2.9)	10.26 (10.2)	31.52 (56.4)	45.31 (48.2)

Figures in brackets are *SD*.

Table 3: Correlations Between Session 1 Private Speech Rates and Task Performance

	Session 1 Performance		*Session 2 Performance*	
	Time	Move-Value	Time	Move-Value
Session 1 PS2	-0.38*	0.17	-0.06	0.09
Session 1 PS3	-0.09	0.34*	-0.03	-0.04

Time = Total Time; Move-Value = Total Move-Value. Correlations incorporate partialling for age and social speech rate (Sessions 1 and 2, d.f.=41). *p<.05, two-tailed.

Table 4: Correlations Between Session 1 Private Speech and Time to Solution

Level of difficulty	PS2	PS3
2-move	0.03	0.02
3-move	-0.18	-0.04
4-move	0.01	-0.09
5-move	-0.12	0.05

A negative correlation represents a positive association between private speech and task performance, as lower times indicate better performance. Correlations include partialling for age and social speech rate (Sessions 1 and 2, d.f. = 41).

Table 5: Concurrent Speech–Outcome Combinations in Relation to Task Difficulty in Session 1, as Percentages of the Total Sample of Children

	Level of difficulty			
	2-move	3-move	4-move	5-move
Relevant/success	17.4	43.5	43.5	26.1
Relevant/failure	8.7	13.0	15.2	28.3
Irrelevant/success	4.3	8.7	4.3	0.0
Irrelevant/failure	2.2	4.3	2.2	2.2
Silence/success	65.2	28.3	30.4	26.1
Silence/failure	2.2	2.2	4.3	17.4

Figure 1: The Tower of London Task

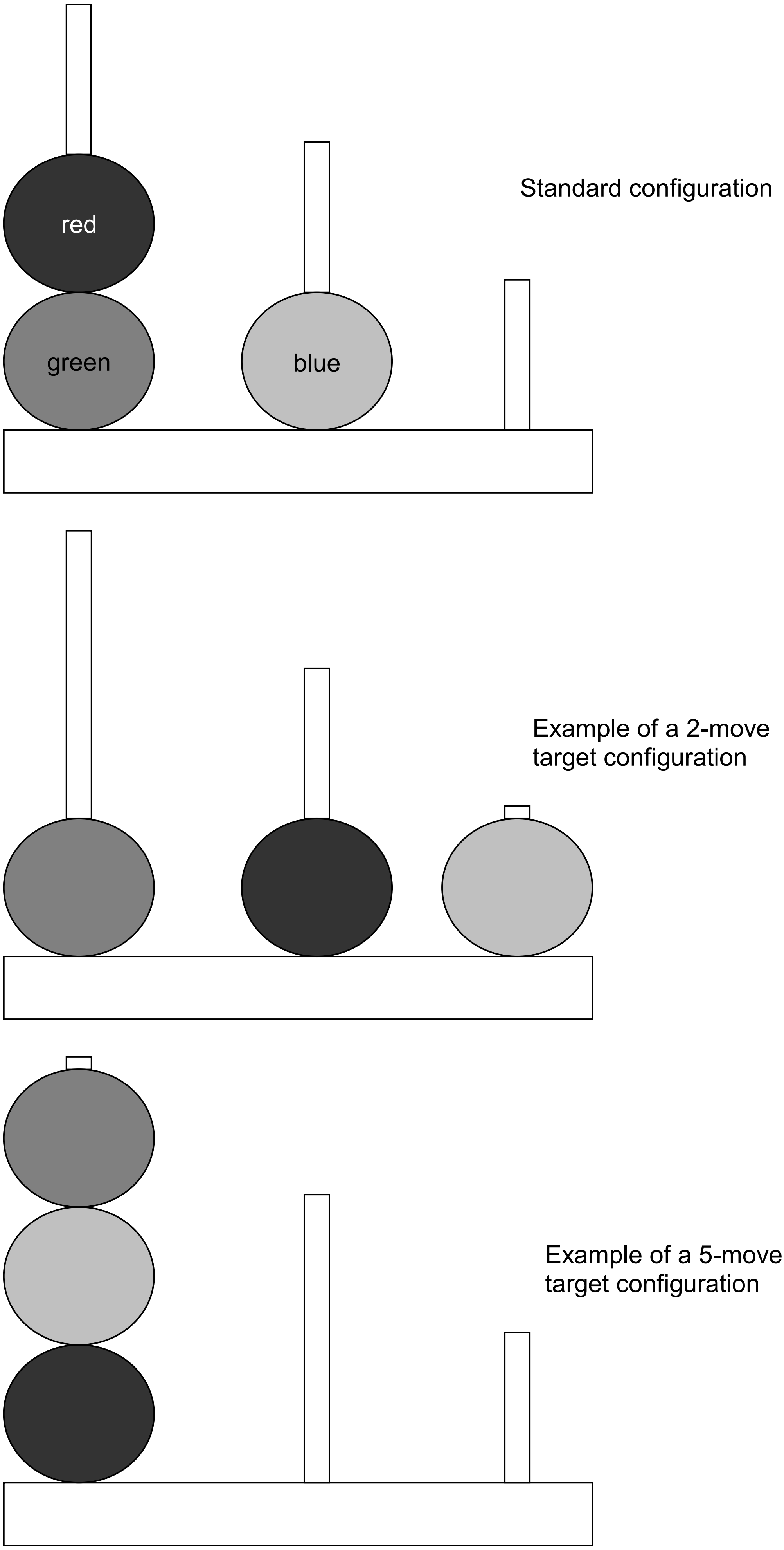

Figure 2: Mean Rates of Speech Sub-Types as a Function of Task Difficulty

(collapsed across Sessions 1 and 2; mean utterances per minute)