Mental Simulations in Comprehension of Direct versus Indirect Speech Quotations

Bo Yao, B.Sc.

Submitted in fulfilment of the requirements for the Degree of Ph.D.

School of Psychology
College of Science and Engineering
University of Glasgow

December 2011

Acknowledgements

I would like to express my greatest gratitude to my supervisors Dr Christoph Scheepers and Prof Pascal Belin for their invaluable guidance and support throughout the time of my Ph.D. research.

Special thanks to Dr Christoph Scheepers who took me on board despite my Biology background. Your saying, 'man with motivation can achieve anything' has inspired me to accomplish my Ph.D. adventure, and has encouraged me to pursue an independent academic career.

I am also grateful to Anna Murray who helped with stimulus preparation and data collection for some part of this research, to Rebecca Watson for her kind tutorials on fMRI data analysis, and to Dr Sara Sereno and Prof Paddy O'Donnell for their kind support and inspiration.

Finally, I would like to thank my parents for their unconditional love and support, and my uncle for guiding me to be a confident scientist and a respectable person.

Author’s declaration

I declare that this thesis is my own work carried out under normal terms of supervision.

Abstract

In human communication, direct speech (e.g., *Mary said: 'I'm hungry'*) coincides with vivid paralinguistic demonstrations of the reported speech acts whereas indirect speech (e.g., *Mary said [that] she was hungry*) provides mere descriptions of what was said. Hence, direct speech is usually more vivid and perceptually engaging than indirect speech. This thesis explores how this vividness distinction between the two reporting styles underlies language comprehension. Using functional magnetic resonance imaging (fMRI), we found that in both silent reading and listening, direct speech elicited higher brain activity in the voice-selective areas of the auditory cortex than indirect speech, consistent with the intuition of an 'inner voice' experience during comprehension of direct speech. In the follow-up behavioural investigations, we demonstrated that this 'inner voice' experience could be characterised in terms of modulations of speaking rate, reflected in both behavioural articulation (oral reading) and eye-movement patterns (silent reading). Moreover, we observed context-concordant modulations of pitch and loudness in oral reading but not straightforwardly in silent reading. Finally, we obtained preliminary results which show that in addition to reported speakers' voices, their facial expressions may also be encoded in silent reading of direct speech but not indirect speech.

The results show that individuals are more likely to mentally simulate or imagine reported speakers' voices and perhaps also their facial expressions during comprehension of direct as opposed to indirect speech, indicating a more vivid representation of the former. The findings are in line with the demonstration hypothesis of direct speech (Clark & Gerrig, 1990) and the embodied theories of language comprehension (e.g., Barsalou, 1999; Zwaan, 2004), suggesting that sensory experiences with pragmatically distinct reporting styles underlie language comprehension.

Publications obtained during the course of the study, of which, materials (or part) will be presented in the thesis:

Journal papers:

- **Yao, B.**, Belin, P., & Scheepers, C. (2011). Silent reading of direct versus indirect speech activates voice-selective areas in the auditory cortex. *Journal of Cognitive Neuroscience, 23*(10), 3146-3152. Presented in Chapter 2.
- **Yao, B.**, & Scheepers, C. (2011). Contextual modulation of reading rate for direct versus indirect speech quotations. *Cognition, 121*(3), 447-453. Presented in Chapter 6.
- **Yao, B.**, Belin, P., & Scheepers, C. (under review). Brain 'talks over' boring quotes: Top-down activation of voice-selective areas while listening to monotonous direct speech quotations. *NeuroImage*. Presented in Chapter 3.

Conference and workshop presentations:

- **Yao, B.**, Belin, P., & Scheepers, C. (June 2011). Activation of voice-selective brain areas in response to written and auditory quotations. *Poster presentation at 17th Annual Meeting of the Organization on Human Brain Mapping (HBM)*, Quebec City, Canada.
- **Yao, B.**, Belin, P., & Scheepers, C. (April 2011). *Poster presentation at* Activation of voice-selective brain areas in response to written and auditory quotations. *International Workshop on Voice and Speech Processing in Social Interations (VSPSI),* Glasgow, UK.
- **Yao, B.**, & Scheepers, C. (April 2011). Spontaneous voice acting during reading aloud of direct versus indirect speech. *Poster presentation at International Workshop on Voice and Speech Processing in Social Interations (VSPSI),* Glasgow, UK.
- **Yao, B.**, Belin, P., & Scheepers, C. (March 2011). Activation of voice-selective brain areas in response to written and auditory quotations. *Poster presentation at 24th Annual Conference on Human Sentence Processing (CUNY),* Stanford, USA.
- **Yao, B.**, & Scheepers, C. (March 2011). Spontaneous voice acting during reading aloud of direct versus indirect speech. *Poster presentation at 24th Annual Conference on Human Sentence Processing (CUNY),* Stanford, USA.
- **Yao, B.**, & Scheepers, C. (March 2011). Contextual modulation of reading rate in oral and silent reading of direct versus indirect speech. *Poster presentation at 24th Annual Conference on Human Sentence Processing (CUNY)*, Stanford, USA.
- **Yao, B.**, Belin, P., & Scheepers, C. (January 2011). Activation of voice-selective brain areas in response to written and auditory quotations. *Poster presentation at Language in Glasgow (LanG)*, Glasgow, UK.
- **Yao, B.**, Belin, P., & Scheepers, C. (September 2010). Top-down activation of voice areas in the auditory cortex during silent reading of direct vs. indirect speech. *Selected talk at 16th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP)*, York, UK.

Table of Contents

Abstract 4

Table of Contents 6

Chapter 1 Thesis introduction 11

1.1 Direct and indirect speech and their pragmatic functions 11

1.2 The vividness distinction between direct and indirect speech in verbal communication 13

1.3 The representational consequences of the vividness distinction in language comprehension 15

1.4 Thesis hypothesis and motivations 19

1.5 Thesis overview 23

Part 1: Differential brain activity in voice-selective areas during comprehension of direct versus indirect speech 24

Part 2: The acoustic aspects of the 'inner voice' in comprehension of direct versus indirect speech 25

Part 3: The visual aspect of the direct-speech representations: Mental simulations of facial expressions? 28

PART 1: Differential Brain Activity in Voice-Selective Areas during Comprehension of Direct versus Indirect Speech 30

Chapter 2 Brain 'hears' voices when reading direct quotations: Silent reading of direct versus indirect speech activates the human voice-selective areas (Experiment 1) 31

2.1 Introduction 31

2.2 Method 33

2.2.1 Participants 33

2.2.2 Stimuli and design 33

2.2.3 Procedure 34

2.2.4 MRI acquisition....36

2.2.5 Data analysis....37

2.3 Results....38

2.4 Supplemental analyses....39

2.5 Discussion....42

Chapter 3 Brain 'talks over' boring quotes: Top-down activations of voice-selective areas when listening to monotonous direct speech utterances (Experiment 2)....46

3.1 Introduction....46

3.2 Method....48

3.2.1 Participants....48

3.2.2 Stimuli....48

3.2.3 Procedure....49

3.2.4 MRI acquisition....50

3.2.5 Data analysis....50

3.3 Results and discussion....55

3.3.1 Whole brain and ROI analyses....55

3.3.2 Parametric modulation analyses....57

3.3.3 Comparing brain activation patterns between studies....64

3.4 Summary and general discussion....65

PART 2: The Acoustic Aspects of the 'Inner Voice' in Comprehension of Direct versus Indirect Speech....69

Chapter 4 Contextual modulation of pitch representations in oral and silent reading of direct versus indirect speech....70

4.1 Introduction....70

4.2 Experiment 3 77
4.2.1 Participants 77
4.2.2 Stimuli and design 77
4.2.3 Procedure 79
4.2.4 Data analyses 79
4.2.5 Results and discussion 81
4.3 Experiments 4, 5 and 6 88
4.3.1 Participants 88
4.3.2 Stimuli and design 88
4.3.3 Procedure 91
4.3.4 Results and discussion 92
4.4 Summary and general discussion 96
Chapter 5 Contextual modulations of loudness representations in oral and silent reading of direct versus indirect speech 99
5.1 Introduction 99
5.2 Experiment 7 101
5.2.1 Participants 101
5.2.2 Stimuli and design 101
5.2.3 Procedure 102
5.2.4 Results and discussion 102
5.3 Experiment 8 104
5.3.1 Participants 104

5.3.2 Stimuli and design....104
5.3.3 Procedure....106
5.3.4 Results and discussion....107
5.4 General discussion....110
Chapter 6 Contextual modulation of reading rates for direct versus indirect speech quotations....112
6.1 Introduction....112
6.2 Experiment 9....113
6.2.1 Participants....113
6.2.2 Stimuli and design....113
6.2.3 Procedure....114
6.2.4 Results and discussion....115
6.3 Experiment 10....116
6.3.1 Participants....116
6.3.2 Stimuli and design....116
6.3.3 Procedure....116
6.3.4 Results and discussion....117
6.4 Summary and general discussion....121
PART 3: The Visual Aspect of the Direct-Speech Representations: Mental Simulations of Facial Expressions?....124
Chapter 7 Are facial expressions also encoded in silent reading of direct speech? (Experiment 11)....125
7.1 Introduction....125
7.2 Participants....128

7.3 Stimuli and design 128

7.4 Procedure 130

7.5 Results and discussion 131

Chapter 8 Conclusions and closing remarks 136

Appendices 141

References 195

Chapter 1 Thesis introduction

1.1 Direct and indirect speech and their pragmatic functions

Direct speech (e.g., *Mary said: 'I'm hungry'*) and indirect speech (e.g., *Mary said [that] she was hungry*) are two important reporting styles in human communication (Coulmas, 1986). Direct speech is commonly used when the reported utterance is framed in the reported speaker's words (e.g., *Mary*'s words) whereas indirect speech is applied when the reported utterance is paraphrased in the reporter's words (e.g., my words).

Direct speech and indirect speech have distinct pragmatic functions. The more traditional view (e.g., Leech & Short, 1981; Thomson & Martinet, 1986) argues that direct speech is a *verbatim* quotation of the reported utterance whereas indirect speech is a *paraphrased* description of the reported utterance. While such a hypothesis sufficiently interprets the usage of indirect speech, it appears to be inadequate in explaining certain phenomena of direct speech usage.

First of all, a verbatim hypothesis implies that a direct speech quotation can only be made upon an *existing* original utterance. However, there are cases in which *no* original utterance *has ever existed* to be quoted. For example, people's thoughts (e.g., *Mary thought: 'I'm hungry'*) have never been uttered into words; therefore the content of thoughts can only be inferred rather than be 'quoted' *verbatim*. Moreover, utterances in fiction do not exist in the real world and it would be illogical to quote Gandalf (i.e., a wizard in *the Lord of the Rings*) word-for-word.

Secondly, a verbatim hypothesis assumes that the quoted original utterance has to be a linguistic statement uttered by a human being such that it can be quoted word-for-word. But in some cases, the original utterances are non-linguistic vocalisations which cannot be understood without imaginary interpretations. For instance, *My kitten was like: 'Meow.. I want milk, Mummy!'* is believed to be a direct quotation of a cat's meowing. Since kittens do not speak English, there could have been no linguistic information available to be quoted. Thus, the *'I want milk'* quotation can only be an *inferred* or *imagined* linguistic interpretation rather than a *verbatim* quotation.

Finally, a verbatim hypothesis also overestimates our memory capacity. Most of the time, we simply cannot remember every single word that has been said. Empirical research shows that it is surprisingly difficult for participants to recall an utterance word-for-word even after just a few seconds (e.g., Anderson, 1974; Hjelmquist & Gidlund, 1985; Jarvella, 1971; Lehrer, 1989; Sachs, 1967; Stafford, Burggraf, & Sharkey, 1987). Moreover, certain expressive elements allowed in direct speech such as 'so-and-so', 'blah-blah-blah' and 'this and that and everything else' clearly suggest missing out of the original words. For example, consider the following sentence *Then we come back to me in the studio and I say, 'The film you have just seen was an attempt to blah blah blah.'* (British National Corpus, G12, 731). In this context, 'blah blah blah' was not the exact words the reported speaker said. Instead, it was intended to represent 'some cliché comments about a film' which the reporter – who was in fact the reported speaker himself – simply could not remember or did not even bother to report word-for-word. Such an intended 'unfaithfulness' in direct speech reporting is clearly incongruent with the *verbatim* hypothesis.

In contrast, a recent proposal by Clark and Gerrig (1990) adequately accommodates the linguistic phenomena that have been discussed above, namely, direct quotations of 'never-existed', non-linguistic and omitted/forgotten utterances. They argue that direct speech, instead of being a verbatim quotation of the reported utterance[1], provides *non-serious* and *selective demonstrations* of the reported speech acts. For example, when the reporter says *Mary said: 'I'm hungry!'*, he/she means that the reported speaker *Mary* is hungry. In this sense, *'I'm hungry!'* is not a serious statement that the reporter meaning him/herself is hungry but a demonstration of the reported speech utterance. Moreover, the reporter may *selectively* demonstrate certain aspects of the original utterance such as dramatizing *her* vocal features while omitting what *she* actually said. As *demonstration*, direct speech does not have to be a *verbatim* quotation of the reported utterance; in some circumstances, it can simply be 'constructed' regardless of the original utterance's existence or nature (Clark & Gerrig, 1990; Tannen, 1986, 1989). Thus, even though the original utterances have never existed, are non-linguistic or are omitted or forgotten, they can still be 'constructed' to

[1] Note: the *demonstration* account is not *'anti-verbatim'*; it actually acknowledges that at least in certain institutional settings, such as in newspapers, law courts or academic contexts, verbatim quotations are conventionally encouraged (cf. Clark & Gerrig, 1990; Short, Semino, & Wynne, 2002).

depict what the reported situations *may have been like* or what the reported speakers *may have said*.

More importantly, Clark and Gerrig's proposal recognises that in speech reporting, paralinguistic information such as speakers' voices, facial expressions and body languages conveys important pragmatic meanings which linguistic information alone cannot effectively communicate. As a result, the core of Clark and Gerrig's (1990) proposal lies in its emphasis on the paralinguistic forms (i.e., how they are reported) rather than the linguistic content (i.e., what was reported) of direct speech reporting. They argue that the key distinction between direct speech and indirect speech lies in the stylistic features of the two reporting styles: Direct speech is an important stylistic device for enlivening stories; it provides vivid demonstrations of the reported speech acts, enabling listeners to experience what it would be like to see, hear or feel what the reported speaker did. By contrast, indirect speech provides mere descriptions of what was said, without depicting the paralinguistic information that was communicated in the reported speech acts. This pragmatic distinction predicts that direct speech is perceptually more vivid and engaging than indirect speech.

1.2 The vividness distinction between direct and indirect speech in verbal communication

Indeed, linguists have observed that direct speech usually coincides with more vivid depictions of the reported speaker's voice, facial expressions and gestures as compared to indirect speech (e.g., Chafe, 1982; Clark & Gerrig, 1990; Labov, 1972; Schiffrin, 1981; Tannen, 1982, 1986, 1989). Moreover, in English, the more frequent use of *be-like*, as a quoting device for direct speech (e.g., *Mary <u>was like</u>, 'I'm so hungry'*) but not indirect speech, also indicates a rising communicative demand for paralinguistic demonstrations for the former. In recent years, *be-like* has become more frequently used and has almost replaced the conventional *said* as the alternative quotative of choice for direct speech among younger generations (Blyth, Recktenwald, & Wang, 1990; Cukor-Avila, 2002; Ferrara & Bell, 1995; Tagliamonte & D'Arcy, 2004; Tree & Tomlinson, 2008). Unlike *said* which can only express something 'in words', *be-like* can be used to describe an event, feeling, or person using both linguistic and paralinguistic information, especially when it is difficult to describe in words. In this sense, the increasing preference for *be-like* direct

speech may indicate a rising trend in using paralinguistic depictions (which are hard to express in words) when reporting direct speech. This trend is consistent with the *demonstration* pragmatic function of direct speech.

In addition to the observational evidence, Clark and Gerrig's *demonstration* theory has also been experimentally tested under laboratory environments. In Wade and Clark's (1993) study, participants first watched short videotaped conversations, which they then had to orally report to a naïve listener who was seated on the opposite side of an opaque screen. Importantly, the narrators were either instructed to amuse their listeners (being entertaining) or to recount the conversations as accurately as possible (being accurate). Analyses of the participants' speech recordings indicated that narrators who were instructed to be entertaining produced a greater proportion of direct-speech quotations, whereas narrators who were instructed to be accurate produced a nearly even distribution of direct and indirect speech quotations. In other words, language producers are more likely to employ direct-speech quotations when communicative demands (such as the instruction to be entertaining) encourage the use of a more vivid, demonstrative reporting style. A longitudinal study by Goodell and Sachs (1992) indicates that the pragmatic distinction between direct and indirect speech is learned over many years and is mastered in adulthood. They showed that 6-year-olds, although they may have acquired the knowledge of indirect speech, were more likely to use direct speech than adults when retelling conversational narratives. However, they used direct quotations one after another, in a more 'self-centred' manner – in the sense that they did not consider which reporting style to use in order to make listeners understand the story more effectively. Using direct speech simply relieved them from a heavy processing load in summarising the conversationally conveyed information (which requires using indirect speech). Adults, on the other hand, used indirect speech to summarise information, and saved direct speech to depict 'highlights' of the story. Their retelling was more 'listener-centred' – in the sense that it attempted to convey information effectively, taking the listener's vantage point into account.

In sum, both the observational and the empirical evidence support Clark and Gerrig's demonstration theory of direct speech. They show that in verbal communication, direct speech is used to provide vivid paralinguistic demonstrations of the reported speech acts and that it is more vivid and engaging than indirect speech which provides mere descriptions of what was said.

1.3 The representational consequences of the vividness distinction in language comprehension

The vividness distinction between direct and indirect speech is obvious during *verbal communication* which is rich in paralinguistic information. Our research question is: How are direct speech and indirect speech mentally represented during *language comprehension*, e.g., silent reading of text, where *little* paralinguistic information is available?

Classic proposals such as the 'amodal symbol system' argue that language is represented in arbitrary, amodal symbols that are independent from the brain's modal systems for perception, action and introspection (e.g., Chomsky, 1957; Fodor, 1975; Pylyshyn, 1984). For example, the sentence *Bo opened the book* can be converted into a propositional representation (e.g., Kintsch, 1998; Kintsch & Vandijk, 1978) such as [OPENED[BO, BOOK]]. The problem with such proposals is that those amodal, symbolic representations are *arbitrarily* linked to the perceptual states that produce them. Just as the printed word *"book"* has no systematic similarities to books that exist in the physical world, the amodal, symbolic representation for *book* has no systematic, physical similarities to the books that we perceive. Thus, just as the printed words "book" and "dictionary" are not necessarily more similar than the printed words "book" and "hospital", the amodal, symbolic representations for *book* and *dictionary* are not necessarily more similar than the amodal representations for *book* and *hospital*. The lack of grounding in the brain's modal systems for perception, action and introspection therefore introduces the transduction problem: How can amodal symbol systems map modality-specific perceptual states onto amodal symbols? By proposing an amodal, arbitrary representation of language, the amodal symbol systems fail to provide a satisfactory account of such a transduction process. Moreover, the lack of cognitive or neural evidence has also questioned the existence of such a transduction process. A related shortcoming of the amodal symbol systems is the symbol grounding problem (Harnad, 1990; Searle, 1980), i.e., how arbitrary, amodal symbols represent the physical similarities or distinctions of referents in reality. This problem can be illustrated by Searle's (1980) Chinese Room problem: Imagine you are in a foreign country like China and unfortunately you do not speak Chinese. You wish to understand Chinese, so you open up a dictionary, trying to study the meanings of Chinese words. However, these words are also explained in Chinese which are just other amodal, arbitrary symbols to you. As a result, you still do not understand Chinese. Metaphorically,

we can view those Chinese definitions as language representations (meanings of language), and the Chinese words you wish to understand as language itself. If the language representations are arbitrary (e.g., Chinese explanations) and do not represent the physical similarities or distinctions of their referents in the physical world, they would be ineffective in indexing the physical referents (i.e., the meanings) of language (e.g., Chinese words) in the physical world. In addition to the theoretically unresolved problems that have been discussed above, the amodal symbol systems also face a shortage of empirical support. In order to verify the existence of amodal symbol systems, we should observe that the full spectrum of language representations are stored and processed in the same brain area that is independent from the sensori-motor circuits (i.e., the supra-modal language areas should exist). At the behavioural level, we should observe that language comprehension stays independent from (i.e., has no interaction with) concurrent perception, actions or introspection. However, over the last 20 years, there has been little direct empirical evidence which verifies the above predictions. By contrast, researchers have frequently discovered that language representations actually have sensori-motor characteristics, and that language comprehension routinely interacts with concurrent perception or actions (details discussed below). These findings oppose the amodal, symbolic account of language representations and are in favour of a modality-dependent, 'embodied' nature of language representations.

The 'embodied' theories propose that language representations are essentially grounded in our sensory and motor experiences of the environment, the body and the mind, and that language comprehension is also implemented at this level via simulations (i.e., partial re-enactments of perceptual, motor and introspective states acquired during experience with the world, body and mind), making language representations well suited for interacting with perception, actions and introspection (e.g., Barsalou, 1999, 2008; Zwaan, 2004). The grounding of language representations in our brain's modal systems resolves the transduction problem and the symbol grounding problem by making no fundamental distinction between language representations and the corresponding perceptual states and actions[2]. Under an embodied framework, we should observe that language representations

[2] This is not to say that language representations and sensori-motor mental states are exactly the same in every single aspect. The idea is that they are qualitatively equivalent in the sense that they are both modality-

are modality-specific and that that they are stored and implemented in different brain areas that we use for perception, action and introspection. At the behavioural level, we should observe interactions between language comprehension and concurrent perception or action. Indeed, behavioural and neuroimaging studies indicate that language representations are flexible, distributed representations comprised of sensori-motor characters. At the behavioural level, numerous studies show that language comprehension routinely activates sensory and motor representations, and that mental simulation of such modality-dependent representations is an integral part of natural language comprehension (for a review, see Barsalou, 2008). For example, researchers showed that individuals mentally simulate objects' orientations (Stanfield & Zwaan, 2001) and shapes (Zwaan, Stanfield, & Yaxley, 2002) that are linguistically implied in sentences (e.g., *John put the pencil in the cup* implies a vertical rather than a horizontal orientation of the pencil; *The ranger saw the eagle in the sky* implies an eagle with its wings stretched out rather than folded). Immediately after reading such sentences, subjects responded differentially to illustrations of objects depicting congruent or incongruent orientations or shapes. These findings suggest that language representations consist of modality-specific perceptual features (e.g., visual features of orientations and shapes) depending on specific sensory experiences (e.g., an eagle with its wings stretched out) that we acquired in the linguistically described situations (e.g., where we see an eagle flying in the sky). In the motor domain, evidence also indicates that understanding linguistic descriptions of actions recruits motor representations that can be characterised in directional movements (e.g., *Close the drawer* implies a movement away from the body; cf. Glenberg & Kaschak, 2002) or rotational movements (e.g., *Eric turned down the volume* implies an anticlockwise rotation; cf. Zwaan & Taylor, 2006); such motor simulations are found to interfere with individuals' concurrent motor responses that are required by a secondary task during reading. For example, Glenberg & Kaschak asked subjects to make a sentence sensibility judgement after reading sentences that imply either a towards-the-body (e.g., *Open the drawer*) or an away-from-the-body (e.g., *Close the drawer*) directional movements. The subject had to make responses by pressing buttons that were either near or far away from him/herself, thereby making a towards-the-body or an away-from-the-body movement. They found that subjects were faster in making responses when their button-pressing movements were congruent (along the towards-vs.-away dimension) with the movements that were implied

specific. This may include the possibility that they may differ from each other to some extent on certain quantitative dimensions (e.g., amplitude).

in the previously read sentences. The findings suggest that motor representations of directional movements can be activated during language comprehension, and interfere with language comprehenders' concurrent motor movements. At the neural level, researchers attempted to map out the neural activations in relation to concepts of different modalities. For example, neuroscientists observed that comprehension of actions words that are semantically related to different parts of our body (e.g., 'lick', 'pick' and 'kick') elicited early (100-200 ms) neural activations in the corresponding motor and premotor cortices (see Pulvermuller, 2005 for a review): Leg-related words activated brain areas that are overlapping and/or adjacent to brain areas involved in foot movements and similar relationships were also observed between comprehension of arm-related and face-related words and neural activity in finger-movement and tongue-movement brain areas. These observations show that language/conceptual representations are indeed stored in distinct sensory and motor brain areas depending on specific sensory and motor experiences we acquired during language/conceptual acquisition.

With respect to our research question, the amodal vs. embodied views of language representations make distinct predictions regarding how direct speech and indirect speech are mentally represented in language comprehension such as silent reading. If we adopt the amodal view of language representations, we would predict that direct speech (e.g., *Mary said, 'I'm hungry'*) and indirect speech (e.g., *Mary said that she was hungry*) are represented in the same way. For instance, they could be converted into identical propositional representations such as [SAID[MARY, WAS[MARY, HUNGRY]]]. This hypothesis seems to be philosophically implausible - if direct speech and indirect speech are represented in exactly the same way, what is the point of having two of them in written language, simply as the legacy from their verbal siblings? A recent study by Bohan, Sanford, Cochrane, and Sanford (2008) demonstrates that direct speech enhances the depth of processing in silent reading as compared with indirect speech. They employed a change detection paradigm: They presented a sentence twice but one critical word was changed to a semantically-similar word during the second presentation (e.g., *I'll just have to wait outside for my* ***flatmate*** → ***roommate*** *to let me in*). They found that readers were more likely to detect such subtle changes when the sentence was presented in a direct speech quotation (e.g., *She said to her boyfriend: "I'll just have to wait outside for my* ***flatmate*** → ***roommate*** *to let me in."*) as opposed to when it was presented in an indirect speech clause

(e.g., *She said to her boyfriend that she would just have to wait outside for her* ***flatmate*** → ***roommate*** *to let her in.*). Their results suggest that direct and indirect speech are processed differently during language comprehension. Moreover, most of us tend to hear an intuitive 'inner voice'[3] particularly during silent reading of direct speech but not indirect speech. Such differential 'inner voice' experiences in understanding the two reporting styles also suggest that they may be mentally represented in distinct manners, at least within the auditory domain. In fact, this voice-related representational distinction between direct and indirect speech in silent reading is naturally predicted by the contrasting, embodied views of language representations. In line with these proposals, the mental representations of direct speech and indirect speech should be grounded in the distinct sensori-motor experiences we acquired when using or perceiving them during verbal communication. In other words, the vividness distinction between direct and indirect speech in verbal communication (i.e., direct speech is more vivid than indirect speech) should underlie language comprehension of written direct and indirect speech: Comprehension of the former should be grounded in mental simulations (re-enactments) of more vivid demonstrations which consist of, e.g., depictions of reported speakers' voices; this could explain our intuition of hearing an 'inner voice' during silent reading of direct speech. By contrast, comprehension of indirect speech should be grounded in mental simulations of descriptions which are less likely to contain reported speakers' voices.

1.4 Thesis hypothesis and motivations

Surprisingly, although naturally predicted by embodied theories of language comprehension, there has been little objectively measurable confirmation of this 'inner voice' experience during comprehension of direct speech to date.

Previous studies have shown that processing of sound-related concepts recruits our auditory system (e.g., Kellenbach, Brett, & Patterson, 2001; Pecher, Zeelenberg, & Barsalou, 2003). For instance, Pecher et al. (2003) found that verifying a property in the

[3] The 'inner voice' here refers to re-enactment of higher-level, supra-segmental acoustic features (e.g., Abramson, 2007; Alexander & Nygaard, 2008; Kosslyn & Matt, 1977; Kurby, Magliano, & Rapp, 2009), rather than to lower-level phonological/prosodic encoding (e.g., Abramson & Goldinger, 1997; Ashby & Clifton, 2005; Ashby & Rayner, 2004) which presumably occurs during silent reading of both direct and indirect speech.

auditory modality (e.g., *BLENDER-loud*) became slower after verifying a property in a different modality (e.g., *CRANBERRIES-tart*) than after verifying a property in the same modality (e.g., *LEAVES-rustling*). These findings suggest that switching modalities in conceptual processing introduces a cost, analogously to the cost of switching modalities in perception. Although their work did not focus on auditory simulations *per se*, it demonstrates mental simulations in the auditory modality. In the same vein, Lynott and Connell (2009) collected ratings to describe a set of 423 adjectives along the five perceptual modalities (i.e., visual, haptic, auditory, olfactory and gustatory). Their data indicate that sound-related object properties (e.g., *beeping* and *meowing*) are perceived predominantly through the auditory sense. Using simultaneous recording of electroencelphalography (EEG) and fMRI, Kiefer, Sim, Herrnberger, Grothe, and Hoenig (2008) demonstrated that within just 150 ms, recognition of words denoting sound-related objects (e.g., "telephone") activates posterior superior and middle temporal gyri (pSTG/MTG) that are also activated during sound perception. Their results provide neural evidence for mental simulations of sounds during conceptual processing. These studies, although not directly related to the 'inner voice' phenomenon, suggest the possibility that language comprehension of speech-related information may recruit the auditory system and result in an 'inner voice' experience. Moreover, studies on auditory imagery in reading (e.g., Abramson, 2007; Alexander & Nygaard, 2008; Kosslyn & Matt, 1977; Kurby et al., 2009) demonstrate that listening to voice samples incurs an 'inner voice' experience during reading. For example, Kosslyn and Matt (1977) demonstrated that readers recode written material into a format analogous to that of the actual speech of the writer after listening to an audio sample of the writer's voice. They showed that readers read faster when the written material was believed to be authored by a fast-speaking person and read more slowly when the purported author spoke slowly. These demonstrations of 'inner voice' experiences, although not directly related to the distinction between direct and indirect speech, provide evidence of an 'inner voice' in the context of language comprehension.

Thus, this doctoral research set out to confirm the 'inner voice' experience in comprehension of direct speech by hypothesising that individuals are more likely to mentally simulate the reported speakers' voices (among many other aspects) in

comprehension of direct speech as opposed to indirect speech. This research had the following theoretical and empirical motivations.

First, this doctoral research was motivated to experimentally validate Clark and Gerrig's (1990) demonstration theory from the perspective of language comprehension. Since the formation of the theory, the supporting evidence has predominantly been provided via observational analyses of conversations (e.g., Holt, 1996; Tree & Tomlinson, 2008) and corpora (e.g., Tagliamonte & D'Arcy, 2004), and via experimental testing but only from the perspective of language production (Wade & Clark, 1993). It is still unknown whether the pragmatic distinctions between direct and indirect speech (i.e., demonstrations versus descriptions) are also reflected in language comprehension such as silent reading of written text. If this research empirically verified the intuitive 'inner voice' experience in language comprehension of direct speech versus indirect speech, it would further validate Clark and Gerrig's (1990) proposal from the perspective of language comprehension, suggesting that the direct speech reporting style is *inherently* more vivid than the indirect speech reporting style regardless of language modality.

Second, this doctoral research was motivated to extend the embodied theories of language comprehension in three respects: (1) to explore mental simulations within the auditory perceptual modality which has received less attention in the literature; (2) to explore mental simulations of language pragmatics, specifically in relation to semantically-equivalent but pragmatically-distinct reporting styles; (3) to explore the underlying neural correlates of mental simulations of voices in natural language comprehension, using state-of-the-art neuroimaging methodologies.

First of all, this research was motivated to extend the empirical demonstrations of mental simulations to the auditory perceptual modality and to the sentence/discourse level. During the last decade, the empirical testing of embodied theories has predominantly focused on the visuo-spatial domain (e.g., Connell, 2007; Fincher-Kiefer, 2001; Kaup, Ludtke, & Zwaan, 2006; Richardson & Spivey, 2000; Stanfield & Zwaan, 2001; Zwaan et al., 2002) and on the motor domain (e.g., Glenberg & Kaschak, 2002; Kaschak & Glenberg, 2000; Kaschak et al., 2005; Richardson, Spivey, Barsalou, & McRae, 2003; Zwaan, Madden, Yaxley, & Aveyard, 2004; Zwaan & Taylor, 2006); for a review, see Fischer & Zwaan, 2008). There has been little evidence of auditory simulations during language

comprehension, especially at the sentence/discourse level (studies have only examined mental simulations of sound-related *concept words*, e.g., Kellenbach, Brett, & Patterson, 2001; Kiefer, Sim, Herrnberger, Grothe, & Hoenig, 2008). Such a bias towards the visuo-spatial and motor domains may have been due to the urge to verify the embodied hypothesis of language comprehension rather than to comprehensively explore such simulation phenomena. However, in order to delineate the full spectrum of embodied phenomena in language comprehension, it is important to extend our current scope of empirical testing to other sensory modalities and to the sentence/discourse level which have not been thoroughly investigated in the literature. This research, as an example of such extensions, explored mental simulations of voice representations during natural discourse comprehension of direct and indirect speech.

Secondly, this research was motivated to pioneer mental simulations of pragmatics during language comprehension. The research so far has mostly focused on mental simulations in relation to the semantics of language. For example, Glenberg and Kaschak (2002) examined how sentences which imply different action directions (e.g., 'Close the drawer', which implies action away from the body versus 'Open the drawer', which implies action towards the body) modulate mental simulation of motor representations during silent reading. They found that participants had difficulty making a sensibility judgement requiring a body response in the mismatching direction. However, little is known about how semantically-equivalent but pragmatically-distinct linguistic information is understood during language comprehension. For example, although direct speech *Mary said: 'I'm hungry'* and indirect speech *Mary said that she was hungry* convey virtually the same semantic meaning, the former is likely to coincide with vivid paralinguistic demonstrations of the reported speech act whereas the latter is reported in a less vivid fashion. Thus, by demonstrating how the vividness distinction between direct and indirect speech underlies language comprehension, this doctoral work would open up a new research dimension of the embodiment of language pragmatics.

Finally, given the wide range of neuroimaging facilities that are available within the Centre for Cognitive Neuroimaging at the University of Glasgow, this research was motivated to apply the state-of-the-art neuroimaging techniques in psycholinguistic research. Previous research on mental simulations in language comprehension has mostly employed behavioural paradigms. For example, some employ reaction time-based paradigms which

consist of language comprehension and a secondary recognition task. Using such paradigms, researchers examine how language comprehension affects individuals' response latencies to the relevant probe stimuli (e.g., a probe picture) that were made relevant by the previous text. Although those behavioural methodologies have their merits, they are limited in specifying the underlying neural correlates of mental simulations during natural language comprehension (without a secondary task). In order to fill this gap, this doctoral research combined event-related functional magnetic resonance imaging (fMRI) with eye-tracking to investigate the neural correlates of the hypothesised 'inner voice' experience during comprehension of direct versus indirect speech. This interdisciplinary approach could inspire new applications in psycholinguistic research, advancing our understanding of the underlying architectures and neural circuitry of language comprehension.

1.5 Thesis overview

The following is a condensed overview of the empirical work that is presented in this thesis. The more in-depth discussions regarding the theoretical backgrounds and the interpretations of the results will be presented in the individual chapters. Overall, this thesis is presented in three parts. The first part reports functional magnetic resonance imaging (fMRI) evidence of an 'inner voice' experience in comprehension of direct versus indirect speech. The second part reports behavioural evidence of this 'inner voice' experience, aiming to characterise the underlying mental representations into acoustic parameters of pitch, loudness and rate. The third part of the thesis ventured beyond the 'inner voice' phenomena and explored mental simulations of facial expressions during silent reading of direct versus indirect speech.

Part 1: Differential brain activity in voice-selective areas during comprehension of direct versus indirect speech

Chapter 2 tested our hypothesis by measuring the 'top-down' activations of the auditory cortex[4] during silent reading of direct as opposed to indirect speech. We reason that if individuals are more likely to simulate 'inner voice' representations during silent reading of direct speech rather than indirect speech, the brain areas which are specialised in processing human voices should become more active when reading the former as opposed to the latter. In other words, we should observe higher brain activity within the human voice-selective areas (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000)[5] during silent reading of direct as opposed to indirect speech. Combining functional magnetic resonance imaging (fMRI) with eye-tracking in human volunteers, we show that silent reading of direct as opposed to indirect speech engenders significantly higher brain activation in the voice-selective areas of the right auditory cortex. The findings suggest that readers are indeed more likely to engage in mental simulations (or spontaneous imagery) of reported speakers' voices when reading direct speech as opposed to meaning-equivalent indirect speech, as part of a more vivid representation of the former. The results are discussed in comparison with research on auditory imagery and lip-reading, and were interpreted in line with the demonstration theory of direct speech (Clark & Gerrig, 1990) and the embodied hypothesis of language comprehension (Barsalou, 1999).

Chapter 3 explored the supra-segmental nature of the observed 'inner voice' experience in comprehension of direct rather than indirect speech. We measured the 'top-down' brain activity in the voice-selective areas when listening to monotonous direct speech utterances versus monotonous indirect speech utterances. We reason that if direct speech is grounded in mental simulations of supra-segmental acoustic information, listeners would have to mentally simulate such information when listening to monotonous direct speech in which such information is minimised; in contrast, listeners would be less likely to simulate such supra-segmental acoustic information when listening to monotonous indirect speech as the

[4] 'Top-down' means that the activation of the *auditory* cortex is NOT triggered by 'bottom-up' *auditory* stimulation itself. In our case, it is triggered by higher-level language comprehension.

[5] The human voice-selective brain areas include the superior temporal gyrus/superior temporal sulcus (STG/STS) areas bilaterally. The maximum of sensitivity to vocal sounds was located along the upper bank of the central part of the right STS area.

latter is not grounded in vivid vocal depictions. Thus, we should observe higher 'top-down' brain activity in the voice-selective areas when listening to monotonous direct speech utterances as compared to monotonous indirect speech utterances. Using fMRI, we found that listening to monotonous direct speech utterances indeed elicited significantly higher brain activations in the voice-selective areas of the right auditory cortex than listening to monotonous indirect speech utterances. Follow-up rating studies on contextual congruency revealed that monotonous direct speech utterances were perceived as incongruent with their preceding contexts whereas monotonous indirect speech utterances were perceived as contextually congruent. Parametric modulations showed that part of the observed differential brain activations were significantly accounted for by such contextual congruency, but not by the stimuli's acoustical characteristics or their vividness perceived by listeners. The results indicate that within a given context, listeners routinely expect vivid vocal depictions for direct speech but not for indirect speech; they are more likely to mentally simulate such vivid vocal depictions (i.e., enriched supra-segmental acoustic representations), when listening to monotonous direct speech (contextually incongruent) as opposed to monotonous indirect speech (contextually congruent). The observed brain activations were reconciled with those observed in silent reading (Chapter 2) since they were located in virtually the same brain areas across the two studies. Such a between-study consistency suggests that the 'inner voice' experience we observed in silent reading may have also been supra-segmental in nature (e.g., intonation and emotional prosody). Moreover, the current study also provides novel demonstrations of top-down processing during intelligible speech perception which is modulated by pragmatically distinct linguistic information.

Part 2: The acoustic aspects of the 'inner voice' in comprehension of direct versus indirect speech

Chapters 2 and 3 together provide brain-activation evidence that individuals are indeed more likely to mentally simulate enriched supra-segmental acoustic representations of the reported speakers' voices during language comprehension of direct speech as opposed to indirect speech. **Chapters 4 – 6** attempted to further decompose such representations in the acoustic parameters of pitch, loudness and rate, using behavioural methodologies.

Chapter 4 investigated whether the 'inner voice' in reading direct speech contains pitch representations of the reported speakers' voices. We reasoned that if the 'inner voice' experience in comprehension of direct rather than indirect speech entails pitch information, it may affect the comprehenders' articulation (in modulations of pitch) or voice perception (facilitation on pitch-congruent voice stimuli). First, using an *oral-reading* paradigm, we observed that oral readers spontaneously adjusted their own pitch in accordance with the reported speaker's emotional arousal and gender that were implied in the previously read linguistic context, however only during oral reading of direct speech but not indirect speech. Moreover, we found that the differential pitch modulations during oral reading of direct versus indirect speech were masked when linguistic focus was exerted upon reported speakers' voices; such a 'voice-selective' masking effect indicates the underlying representations of such pitch modulations are likely to be voice-related in nature. The findings suggest that oral readers are more likely to mentally simulate reported speakers' pitch representations during oral reading of direct as opposed to indirect speech. Next, we developed three reaction time-based paradigms to examine mental simulations of pitch during *silent reading* of direct versus indirect speech. We examined the effects of reported speakers' genders on the participants' recognition latencies to spoken words that were recorded from either a same-gender (as the reported speaker) speaker or a different-gender speaker (the stimuli only significantly differed in pitch). Using a serial visual presentation paradigm in which the written critical words (for silent reading) and the spoken probe words (for recognition) were presented simultaneously, we found that the participants responded more slowly to the different-gender spoken words than to the same-gender spoken words, however only during silent reading of direct speech quotations that were reported from female speakers (e.g., *She said: '... '*). The results show that silent readers are more likely to mentally simulate the gender-accordant pitch representations during silent reading of direct speech rather than indirect speech, but only when the reported speakers are female; this discrepancy between male and female reported speakers may be interpreted in the acquired experience that female speakers are generally more emotional and expressive when reporting direct speech during verbal communication (e.g., Kring & Gordon, 1998; Robinson & Johnson, 1997). Although interesting, this interpretation is pending further confirmation with fully-counterbalanced materials and more carefully-manipulated auditory stimuli.

Chapter 5 examined whether the 'inner voice' experience entails reported speakers' loudness information during comprehension of direct versus indirect speech. Similar to our previous investigations on pitch simulations, we examined the effects of the purported loudness simulations on participants' articulation and voice perception. In two experiments, we asked participants to orally (Experiment 7) or silently (Experiment 8, reaction time-based method) read written stories that contained either direct speech or indirect speech quotations. The context preceding those quotations described a situation which implied either a loud-speaking or a quiet-speaking quoted protagonist. We found that this context manipulation modulated the participants' reading loudness during oral reading of direct speech but not indirect speech. The findings suggest that oral readers may mentally simulate loudness representations during oral reading of direct rather than indirect speech. However, in silent reading, the results were not very straightforward. We examined the effects of the contextually-described loudness of the reported speakers on the participants' recognition latencies to spoken words that were either loud or quiet in volume. We found that the loudness (volume) levels of the spoken words did not significantly affect the participants' response times, suggesting that loudness representations may not be encoded in vocal volume during silent reading of direct or indirect speech. Unexpectedly, we found a significant context effect in the indirect speech condition rather than the direct speech condition: the participants responded to the spoken words significantly faster when reading a loud-speaking context as opposed to a quiet-speaking context, but only during silent reading of indirect speech rather than direct speech. This context effect may reflect modulations of response times by the differential emotional intensities that are confounded with reported speakers' loudness; such RT modulations by emotional intensity may be cancelled out during silent reading of direct speech due to the integration of the simulated voice representations with the perceived voices of the spoken words. Reconciled the current findings with the literature, it was speculated that speakers' loudness may not be encoded in vocal volumes, but perhaps in other forms.

Chapter 6 studied whether the 'inner voice' experience represents reported speakers' speaking rates during comprehension of direct versus indirect speech. In two experiments, we asked participants to either orally (Experiment 9) or silently (Experiment 10, eye-tracking) read written stories that contained either a direct speech or an indirect speech quotation. The context preceding those quotations described a situation that implied either

a fast-speaking or a slow-speaking quoted protagonist. We found that this context manipulation affected reading rates (in both oral and silent reading) for direct speech quotations, but not for indirect speech quotations. The findings suggest that readers are more likely to engage in mental simulations of the reported speakers' voices when reading direct as opposed to indirect speech, and that the speaking rate is an important acoustic aspect of such voice representations.

Part 3: The visual aspect of the direct-speech representations: Mental simulations of facial expressions?

We have obtained neuroimaging evidence which suggests that individuals are indeed more likely to mentally simulate enriched supra-segmental acoustic representations of reported speakers' voices during comprehension of direct speech as opposed to indirect speech. We have also obtained behavioural evidence which indicates that the simulated voice representations in reading of direct versus indirect speech can be characterised in modulations of acoustic parameters such as speaking rate.

However, in verbal communication, reported speakers' facial expressions and gestures are usually co-demonstrated with voices during direct speech reporting. To comprehensively examine the grounding of language in such *multi-dimensional* demonstrations of direct speech, **Chapter 7** ventured beyond the investigations of 'inner voice' experience and explored whether reported speakers' facial expressions are mentally simulated during comprehension of direct speech rather than indirect speech. We examined the effects of direct versus indirect speech quotations on the recognition latencies to graphic emoticons which were made relevant by the contexts preceding those quotations. Participants silently read written stories containing either a direct speech or an indirect speech quotation, each followed by a graphic emoticon, and indicate whether the picture was congruent with the story they had just read. The facial expression illustrated in the emoticon either matched or mismatched the expression of the reported speaker implied in the story. Three expressions (*happy*, *surprised* and *angry*) were examined. Because participants frequently confused the *surprised* expression with either the *happy* or *angry* expression, we excluded the *'surprised'*-related items to ensure the validity of the to-be-examined *expression-match* effects. On the remaining *'surprised'*-free trials, we found that responses to the graphic emoticons were facilitated when the illustrated facial expression matched the implied

expression of the reported speaker, but only after silent reading of direct speech. This *expression-match* effect was 'masked' when the linguistic focus was manipulated on reported speakers' facial expressions (i.e., both direct and indirect speech quotations were preceded with explicit descriptions of facial expressions, e.g., *She smiled*). Such a 'face-selective' masking effect suggests that the representational differences underlying this *expression-match* effect may be facial expression-like in nature. Despite the fact that many contrasts did not reach significance due to the exclusion of a considerable number of '*surprised*' trials, the result patterns were highly consistent with our hypothesis. In conjunction with the face-related brain activations that were additionally observed in silent reading of direct rather than indirect speech (Chapter 2), the findings imply that silent readers might indeed simulate reported speakers' facial expressions during silent reading of direct speech more so than during silent reading of indirect speech. Nevertheless, future research is needed to replicate the current results with a larger number of items, and using different, unambiguous expressions.

In sum, my PhD work provides strong support for the intuitive 'inner voice' experience we have when reading direct speech but not indirect speech. It demonstrates that individuals are more likely to mentally simulate enriched supra-segmental acoustic representations of reported speakers' voices during comprehension of direct speech as opposed to indirect speech. Simulated voice representations seemed to be sensitive to modulations of acoustic parameters such as speaking rate but not pitch or loudness. Further investigations, however, may yet reveal differences by modulations of pitch or loudness using more sensitive experimental paradigms. Moreover, preliminary results suggest that facial expressions may also be part of a more vivid, *multi-dimensional* representation of direct speech in language comprehension. The findings support Clark and Gerrig's (1990) demonstration theory of direct speech from the persepective of language comprehension and extend the empirical work on embodied theories of language comprehension (e.g., Barsalou, 1999) to the auditory perceptual modality and the embodiment of language pragmatics, with the application of interdisciplinary methodologies.

PART 1: Differential Brain Activity in Voice-Selective Areas during Comprehension of Direct versus Indirect Speech

Chapter 2 Brain 'hears' voices when reading direct quotations: Silent reading of direct versus indirect speech activates the human voice-selective areas (Experiment 1)

2.1 Introduction

As discussed in Chapter 1 (see 1.2), direct speech (as in *Mary said: 'I'm hungry'*) is assumed to entail vivid demonstrations of the reported speech acts, whereas its counterpart indirect speech (as in *Mary said that she was hungry*) provides mere descriptions of what was said (Clark & Gerrig, 1990). In verbal communication, direct speech demonstrations are usually more vivid and perceptually engaging than indirect speech descriptions. This vividness distinction has been observed and discussed by linguists (e.g., Tannen, 1986, 1989). It has also been shown, for instance, that in reporting previously overheard dialogues, speakers are more likely to employ direct rather than indirect speech when instructed to be entertaining to a listener (Wade & Clark, 1993). In their study, participants first watched short videotaped conversations, which they then had to orally report to a naïve listener who was seated on the opposite side of an opaque screen. Importantly, the narrators were either instructed to amuse their listeners (being entertaining) or to recount the conversations as accurately as possible (being accurate). Analyses of the participants' speech recordings indicated that narrators who were instructed to be entertaining produced a greater proportion of direct speech quotations, whereas narrators who were instructed to be accurate produced a nearly even distribution of direct and indirect speech quotations. In other words, language producers are more likely to employ direct speech quotations when communicative demands (such as the instruction to be entertaining) encourage the use of a more vivid, demonstrative reporting style.

However, little research so far has addressed the question of how the two reporting styles are represented in language comprehension, particularly during silent reading of text where no auditory stimulation, or visual stimulation other than text, is present. While many of us share the intuition of hearing an 'inner voice' during silent reading of direct rather than indirect speech, there has been hardly any objectively measurable confirmation of this experience so far. This is surprising, especially given that recent embodied theories (e.g., Barsalou, 1999, 2008) propose that language comprehenders mentally simulate linguistically described situations based on generalised perceptual experiences they have

acquired in the past. Such theories suggest that, even during silent reading of text, direct speech may be more likely to activate 'audible-speech'-like representations than indirect speech. In other words, readers may be more likely to mentally simulate the reported speaker's voice (or aspects thereof) during silent reading of direct as opposed to indirect speech.

One way to test this hypothesis is using functional magnetic resonance imaging (fMRI) to measure 'top-down' activation of the auditory cortex during silent reading of direct versus indirect speech. The fMRI technique has been developed based on the idea that local cerebral blood flow and blood oxygenation reflects local neural activity (Roy & Sherrington, 1890): It measures changes in local deoxyhemoglobin concentration, Blood-oxygen-level-dependence (BOLD), which is reflected in alterations in signal intensity of magnetic resonance images (Ogawa & Lee, 1990; Ogawa, Lee, Kay, & Tank, 1990; Ogawa, Lee, Nayak, & Glynn, 1990). Brain images characterising changes of BOLD signals thereby indicate modulations of neural activity within certain brain areas as a function of manipulated psychological processes. From the fMRI literature, it is known that certain brain areas in the auditory cortex are selectively sensitive to human voices when stimulated 'bottom-up' via auditory sound clips (Belin et al., 2000). Belin and colleagues found that bilateral brain areas along the upper bank of the superior temporal sulcus (STS) showed greater brain activity when participants listened passively to vocal sounds (e.g., coughing, laughter and speech), than to non-vocal environmental sounds (e.g., telephone ringing). Their findings provide us with clearly defined regions of interest for our investigations of the activation of voice-related mental representations during on-line silent reading. Moreover, studies on non-verbal (Bunzeck, Wuestenberg, Lutz, Heinze, & Jancke, 2005; Yoo, Lee, & Choi, 2001) and verbal (Jancke & Shah, 2004) auditory imagery, as well as experiments on visual speech perception (also known as *lip-reading*) (Calvert et al., 1997; MacSweeney et al., 2000), indicate that areas within the auditory cortex are also prone to 'top-down' activation without external auditory stimulation. Furthermore, Spitsyna and colleagues (2006) showed that neural pathways for silent reading and speech perception converge in the superior temporal sulcus (STS) regions of the auditory cortex, suggesting the possibility that silent reading might recruit some auditory processing. Hence, if readers are more likely to engage in mental simulations of reported speakers' voices during silent reading of direct speech, voice-selective areas in the

auditory cortex (Belin et al., 2000) should display enhanced 'top-down' activation as soon as readers come across a direct speech quotation (as opposed to a meaning-equivalent indirect speech statement) in written text. The following event-related fMRI experiment aimed at testing this hypothesis. Participants silently read a number of short written stories for comprehension while their eye-movements and brain activations were simultaneously monitored.

2.2 Method

2.2.1 Participants

In total, 26 adult participants were recruited and scanned. All of them were native English speakers with normal vision and hearing and having no learning or reading disabilities nor history of neurological or psychiatric disorders. Ten participants had to be excluded from analysis due to either (a) no clear response in the voice localiser task (see below) and/or excessive head-movements during scanning (8 subjects), (b) eye-tracking data loss (1 subject), or (c) less than 70% answering accuracy on comprehension questions (1 subject). Data from the remaining sixteen participants (age 18 – 44 years, six males and ten females) were used for the final analyses. All of them were right-handed except for one female subject.

2.2.2 Stimuli and design

Silent reading stimuli

Ninety short stories with different protagonists (indicated by different names) were prepared as reading materials (Appendix 1). Each story started with two declarative sentences to set up a scenario (e.g., *PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. Ella asked for an extension but Jim looked concerned.*). This was followed by either a direct or an indirect speech quotation (e.g., *He said: 'Hmm, we really need those data in by next month for that conference.*' or *He said that they really needed those data in by next month for that conference.*). The reported quotations in both conditions (underscored in the above examples) were equivalent in terms of linguistic content. Comprehension questions (e.g., *Was Ella Jim's*

PhD student?) were also prepared for 23 stories (ca. 25%) to assess participants' overall comprehension accuracy and to ensure that they read the stories for comprehension.

Two lists of stimuli with counterbalanced item-condition combinations (45 direct and 45 indirect speech trials per list) were constructed. Each item appeared only once per list, but in a different condition across lists. Half of the participants saw presentation list one and the other half presentation list two. The presentation order of the items within each list was randomised for each participant.

Voice localiser stimuli

For the voice localiser session (see 2.2.3 Procedure), we presented blocks of vocal sounds and non-vocal sounds provided by the Voice Neurocognition Laboratory (http://vnl.psy.gla.ac.uk), CCNi, University of Glasgow. These stimuli were the same as those employed in Belin et al. (2000), and comprised both speech (e.g., spoken vowels) and non-speech (e.g., laughing and coughing) vocal sound clips, as well as non-vocal sound clips (e.g., telephone ringing, dog barking and bell chiming). The contrast in brain activity elicited by vocal versus non-vocal sounds reliably localises voice-selective areas of the auditory cortex.

2.2.3 Procedure

Participants were positioned in the scanner, wearing goggles (NordicNeuroLab, Bergen, Norway) for visual presentation and eye-tracking (Viewpoint Eye-Tracker, Arrington Research, Inc.) as well as MRI-compatible, electrostatic headphones (NordicNeuroLab) for noise attenuation during fMRI scanning and for auditory presentation during the voice localiser session (see below). After a brief eye-tracker calibration procedure, the main reading session followed during which the participant's brain was scanned and their eye movements recorded. Participants were instructed to read the text silently and carefully so as to be able to answer comprehension questions which would follow after 25% of the short stories they had read. The stimulus materials were presented using E-prime 2.0 (Psychology Software Tools, Inc., USA). The text was presented in a black 15-point Courier New font on a light grey background. The whole text was centred and wrapped within 75% width of the screen, together making the reading stimuli appear as natural as

possible. Each trial began with the presentation of a fixation cross in the middle of the screen for 2 seconds, followed by the presentation of the text for reading. Each story was presented in a sentence-by-sentence fashion. The presentation duration per sentence display was determined as W × 100 ms + S × 50 ms (where W refers to the number of words and S to the number of syllables per sentence), allowing sufficient time for reading. Mean presentation durations for the final (critical) sentence display were 7514 ms (*SD* = 1412 ms) and 7526 ms (*SD* = 1354 ms) for the direct and indirect speech condition, respectively. About 25% of the text presentations were followed by a comprehension question regarding the content of the preceding story. Each such question appeared in the middle of the screen, prompting a 'yes' or 'no' response which participants could provide by pressing buttons on a response box with their index or middle fingers, respectively. Figure 1A provides a schematic illustration of the experimental trial sequence. The 90 reading trials were evenly interspersed with five 'baseline' trials during which a plain fixation cross was presented in the centre of the screen for 30 seconds.

After the main reading session, an anatomical scan of the participant's brain was performed, followed by a brief, 10-min voice localiser scanning session. During the latter, participants were instructed to close their eyes while listening to 20 8-second blocks of vocal and 20 8-second blocks of non-vocal auditory stimuli presented in an efficiency optimised, pseudo random order along with 20 blocks without stimulation acting as a baseline (Figure 1B; cf. Belin et al., 2000).

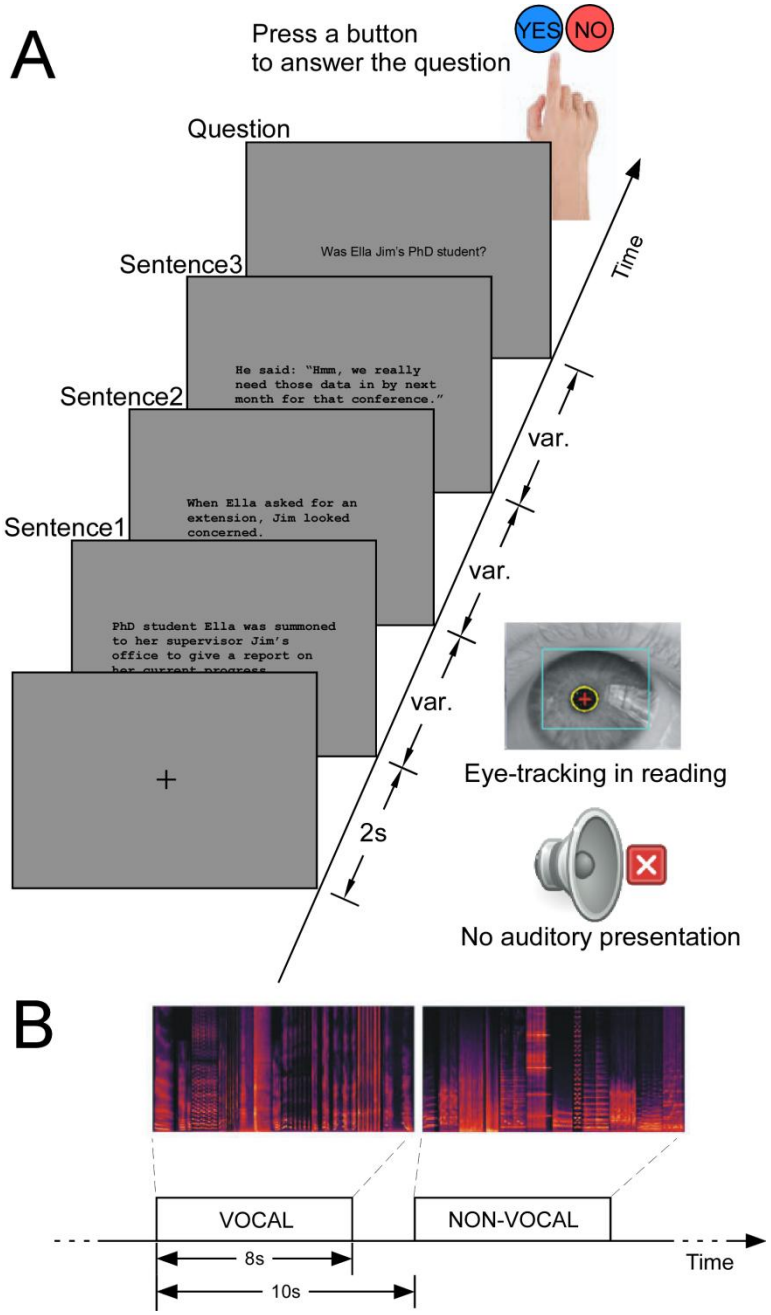


Figure 1 Experimental procedures for (see Experimental Procedures), A, Silent reading task, **B,** Voice localiser task.

2.2.4 MRI acquisition

Scanning was performed on a 3-T Siemens Tim Trio MRI scanner using a 12-channel head coil (Erlangen, Germany). Functional scans (for both the reading session and voice localiser session) were acquired using a T2*-weighted Echo Planar Imaging (EPI) sequence (32 slices acquired in orientation of the Sylvian fissure; TR = 2s; TE = 30 ms; matrix size: 70 × 70; voxel size: 3 × 3 × 3 mm; FOV = 210). T1 whole-brain anatomical scans were obtained using 3D T1-weighted magnetization prepared rapid acquisition

gradient echo (MP-RAGE) sequence (192 axial slices; matrix size: 256 × 256; voxel size: 1 × 1 × 1 mm; FOV = 256). The average scanning time for the whole experiment was around 53 min per participant.

2.2.5 Data analysis

All MRI data were analysed using SPM8 (www.fil.ion.ucl.ac.uk/spm/, University College London). Pre-processing of functional scans included (a) head motion corrections (tri-linear interpolation) whereby scans were realigned to the first volume; (b) co-registration of functional scans to their corresponding individual anatomical scans; (c) segmentation of the co-registered scans; (d) normalisation of functional (3 mm isotropic voxels) and anatomical (1 mm isotropic voxels) data to the Montreal Neurological Institute (MNI) space; (e) smoothing of normalised data (8-mm Gaussian kernel).

fMRI data from the anatomical and the voice localiser scanning sessions were used to determine the voice-selective areas in the auditory cortex of each participant at $p < 0.001$ (Family-Wise Error (FWE)-corrected). The group voice localiser was obtained at $p < 0.001$ (uncorrected, to increase sensitivity against the background of individual differences).

For the reading scanning session, the temporal onset of a critical fMRI-event was defined (via eye-tracking) as the temporal onset of the first fixation in the first continuous reading of the direct or indirect speech quotation in the text; its offset was defined as the temporal offset of the last fixation in the first continuous reading of the direct or indirect speech quotation. Average critical fMRI-event durations amounted to 3118 ms (*SD* = 1175 ms) and 3266 ms (*SD* = 1209 ms), for the direct and indirect speech conditions, respectively. The 148 ms difference in the mean durations is most likely due to the fact that the indirect speech quotations were on average 0.8 words longer than the direct speech quotations. In fact, on a reading time per word measure, there was no appreciable difference between the two conditions (direct speech: 204 ms per word; indirect speech: 203 ms per word; $p > 0.5$ by paired-samples *t*-test). Hence, when minor differences in numbers of words are controlled for, it appears that direct and indirect speech quotations were virtually identical in terms of processing difficulty.

Non-critical readings (of background sentences, comprehension questions, instructions, etc.) and events (button pressing) were specified as corresponding events in the design matrix. The rest consisted of all fixation-cross events (including five 30-second baseline trials and all 2-second pre-trial fixation crosses) and were regarded as baseline. The fMRI data were mapped to the human Colin atlas (sumsdb.wustl.edu/sums/) surface in CARET (brainvis.wustl.edu/wiki/index.php/Caret:About; (D. C. Van Essen, 2002; D.C. Van Essen, Harwell, Hanlon, & Dickson, 2005). The mean beta estimates within ROIs were calculated using SPM toolbox easyROI (www.sbirc.ed.ac.uk/cyril/cp_download.html), and submitted to 2-tailed paired-samples *t*-tests.

2.3 Results

The answering accuracy on the post-trial comprehension questions amounted to 83% (direct speech condition) versus 82% (indirect speech condition). The 1% difference was not significant ($p > 0.4$ by logit binomial GEE; Hardin & Hilbe, 2003).

During critical events (determined via eye-tracking, see 2.2.5 Data analysis), it was found that the direct speech condition was associated with a greater blood-oxygen-level dependent (BOLD) signal in voice-selective brain areas of the right auditory cortex than the indirect speech condition. Although both conditions were more active against the baseline, reading of direct speech elicited significantly greater activation in these areas than reading of indirect speech (see Figure 2C). For individual Regions of Interest (ROIs), the mean between-condition difference amounted to 0.240 ± 0.062 (SE) (2-tailed paired-sample $t(15) = 3.85$, $p = 0.002$); for the group ROI, the between-condition difference was 0.147 ± 0.045 ($t(15) = 3.25$, $p = 0.005$). Three main clusters of enhanced activity for direct as opposed to indirect speech were located in voice-selective areas along the posterior, middle and anterior parts of the right superior temporal sulcus (STS) (Figure 2B). In addition, activation in brain areas other than the auditory cortex was found distributed in the occipital lobes, superior parietal lobules and precuneus (Figure 2A). Although not central to our hypothesis, one might speculate that activation of those areas is part of an enriched multi-sensory mental simulation process for direct speech that also encompasses, for instance, visual aspects of the described situation. As a whole, no region showed an opposite pattern of activity, i.e., direct speech was always associated with a greater BOLD signal than indirect speech. These results support the hypothesis that readers are more

likely to engage in vivid mental simulations of the reported speaker's voice (or aspects thereof) during silent reading of direct speech as opposed to meaning-equivalent indirect speech.

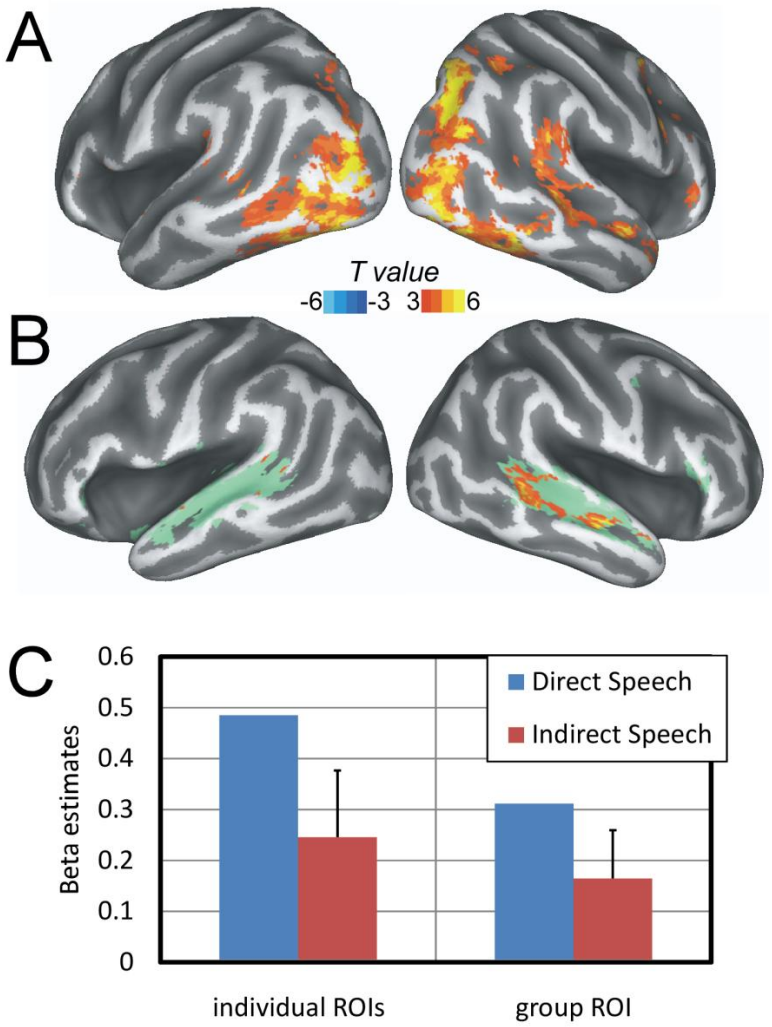


Figure 2 Illustrations of between-condition differences, A, Regions that selectively responded to silent reading of direct as opposed to indirect speech in the whole-brain analysis, **B,** Regions that selectively responded to silent reading of direct as opposed to indirect speech within the sample voice-sensitive areas (green, n = 16), **C,** Mean signal change (against baseline) between conditions in the right voice area Regions of Interest (ROIs), determined individually (left) or using the sample average (right). The single error bar in each panel refers to the 95% CI for the between-condition difference.

2.4 Supplemental analyses

One might point out that our direct versus indirect speech materials sometimes differed in the use of exclamations (e.g., *She said: 'My life has been amazing!...'*), such that 39 of our 90 direct speech items, but none of our 90 indirect speech items, employed emphasis-enhancing exclamation marks. Moreover, the direct and indirect speech conditions were

not always parallel in terms of syntactic coordination (e.g., main clause followed by another main clause) versus subordination (e.g., main clause followed by a subordinate clause). Upon inspection of our stimuli, we found that the latter strictly applies to 11 items. Of the remaining 79 (syntactically more or less parallel) items, 43 contained direct and indirect speech sentences that can be regarded as fully equivalent in terms of syntactic structure.

We addressed these potential concerns as follows. There were 34 items in our set of stimuli (Appendix 1) that neither contained exclamations in the direct speech condition, nor implied any substantial syntactic differences between direct and indirect speech (i.e., from the 43 items where direct and indirect speech shared the same syntactic structure, we excluded 9 items that contained exclamations). The item numbers of these 34 items can be found below. We re-analysed our data (using the same threshold criteria as before) considering only this restricted set of items, which naturally implied a considerable reduction in statistical power due to fewer data points per participant (one third of all the data available). Nevertheless, we found that the previously established effect (over the whole set of items) did not completely disappear when only those 34 items were considered: direct speech was still associated with more activation in the regions of interest than indirect speech. For individual Regions of Interest (ROIs), the mean between-condition difference amounted to 0.162 ± 0.077 (SE) (2-tailed paired-sample $t(15) = 2.093$, $p = 0.054$); for the group ROI, the between-condition difference was 0.087 ± 0.073 (the latter was not significant, but still pointing in the right direction). We then ran analyses comparing these 34 'uncontroversial' items with the remaining 56 'controversial' ones, constituting the factor item group in a 2 [*Item group*: controversial versus uncontroversial] × 2 [*Reporting style*: direct versus indirect] within-subjects ANOVA. Consistent with our original *t*-test analyses over the whole set of items, we found a significant main effect of *reporting style* (individual ROIs: $F(1,15) = 11.602$; $p < .005$; group ROI: $F(1,15) = 6.983$; $p < .02$). However, the main effect of *item group* was far from significant (*F*-values < 1), and so was the *item group* by *reporting style* interaction (*F*-values < 1.3) – the latter would be expected to approach significance if exclamations and/or syntactic complexity differences within the 56 controversial items were mainly responsible for the *reporting style* effect over the whole set of items. Taken together, these analyses suggest that our main results cannot be reduced to differences in emphasis and/or syntactic structure

between the direct and indirect speech conditions. Finally, the fact that we did not find any substantial differences in reading time per word or comprehension accuracy between the two conditions (see 2.2.5 Data analysis & 2.3 Results) may also be taken as an argument against syntactic complexity as a potential confound – at the very least, it suggests that potential syntactic complexity differences between direct and indirect speech were not processing-relevant.

*The 34 items chosen for re-analysis were (item numbers in Appendix 1): 2, 3, 5, 7, 8, 10, 11, 16, 22, 25, 27, 28, 29, 30, 31, 32, 36, 38, 39, 40, 41, 48, 52, 53, 54, 55, 60, 66, 69, 70, 72, 78, 80, 88.

However, even in the subset of 34 items, many direct speech utterances used interjections such as "Oh!", "Hmm…" and "Ah!" which did not appear in the corresponding indirect speech utterances. One concern is that most of the interjections, like exclamation marks, tend to be emotionally arousing and therefore may have confounded the higher brain activations in the right STS area that we observed in the direct speech condition. Admittedly, we cannot rule out that direct speech tends to be associated with increased emotionality. In fact, the reason why direct speech is more vivid or expressive than indirect speech is that when using direct speech, reporters demonstrate or depict the reported speech act, which may also include a demonstration of the quoted speaker's emotional state. We assume that emotionality and voice cannot be ultimately disentangled in speech because voice (just like facial expression) is an important "vehicle" of the reported speaker's emotion, and emotion (or supra-segmental acoustic information that encodes emotion in a voice) may be an important aspect of what is mentally simulated in our experiment. For this reason, it would be implausible to claim that mental simulations of voice during silent reading of direct speech are emotion-free. That said, it is also very unlikely that our data reflect emotionality 'per se'. First, a recent finding from the literature suggests that modality-unspecific emotions are represented in the medial prefrontal cortex (MPFC) and left pSTS (Peelen, Atkinson, & Vuilleumier, 2010) rather than the right STS. Second, our analyses explicitly targeted regions in the brain that are sensitive to voice, using a voice-localiser task that identified regions of interest via bottom-up stimulation. Therefore we are not convinced that increased emotionality of direct speech "on its own" would provide a feasible alternative interpretation of our data – at the very least, voice must be taken into consideration as well. Finally, we are leaning towards

discussing emotional prosody (Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006) as a potential candidate for the kinds of voice-related representations that may be activated during reading of direct as opposed to indirect speech (see Discussion below). This interpretation thereby adequately integrates enhanced voice simulations with increased emotionality in the direct speech condition.

2.5 Discussion

Overall, our results lent objective empirical support to the intuitive experience of an 'inner voice' during silent reading of written text, particularly during silent reading of direct speech. Specifically, our experiment showed that voice-selective areas in the auditory cortex become more active during silent reading of direct speech as opposed to meaning-equivalent indirect speech. This finding could not plausibly be attributed to differences in processing difficulty or comprehension performance because there was no significant difference between the two conditions in terms of reading time per word or question-answering accuracy. Other factors, the use of direct speech exclamations and potential syntactic complexity differences in some of our direct versus indirect speech item pairs, also failed to conclusively account for the observed brain activations.

Previous behavioural studies on auditory imagery (Abramson, 2007; Alexander & Nygaard, 2008; Kurby, Magliano, & Rapp, 2009) also observed the sort of 'inner voice' during silent reading. However, these studies predominantly used such an experimental setup in which participants were presented with particular voices before the actual reading trials, followed by explicit imagery instructions combined with cues to the identity of the to-be-imagined speaker within the actual reading materials. This type of experimental manipulation (pre-exposure to specific voices and/or explicit imagery cues and instructions) arguably encourages participants to imagine speaker-specific voices during silent reading. Hence it is limited in determining whether auditory imagery would also occur as a result of 'normal' language comprehension.

The current experiment focused on (1) whether silent readers activate voice-related representations even without being encouraged to do so (reading for comprehension only) and (2) whether these representations are modulated by different linguistic reporting styles (direct versus indirect speech). A setup was used whereby no particular speaker was

introduced to participants before reading the text passages and in which participants were in no way instructed to imagine voices (all they were asked to do was to read short stories and to answer questions about those stories). Each story contained a unique set of fictitious names and characters such that participants were not (or at least not obviously) led towards imagining any concrete, familiar voices. Combined with the fact that post-reading comprehension questions focused on semantic content (distracting participants' attention away from the direct versus indirect speech manipulation), it seems unlikely that our participants felt encouraged to imagine specific voices during reading. Nonetheless, they still appeared to automatically activate voice-related perceptual representations, particularly in response to reading direct speech quotations.

The clearly right-lateralised brain activation pattern observed in our study also suggests *spontaneous* (in contrast to *encouraged*) activations of voice-related representations during silent reading of direct as opposed to indirect speech. Previous studies on auditory imagery and on lip-reading (Bunzeck et al., 2005; Calvert et al., 1997; Jancke & Shah, 2004; MacSweeney et al., 2000; Shergill et al., 2001; Yoo et al., 2001) mostly observed bilateral (or sometimes left-dominant) top-down activation patterns within the auditory cortex. These studies employed tasks with overt meta-cognitive judgments (explicit imagery with or without a visual cue, shadowing, instructed rehearsing, etc.) which might recruit the left hemisphere more than would be the case in a less explicit experimental setting. The present experiment, by contrast, indicated a clearly right-lateralised locus of effect for the direct versus indirect speech comparison (Figure 2B) using a task that did not involve any meta-cognitive judgments.

To account for this kind of spontaneous auditory imagery during silent reading, as well as its modulation through linguistic reporting style (direct versus indirect speech), one might adopt the notion of *mental simulation* in language comprehension, as proposed by embodied cognition theories (e.g., Barsalou, 1999, 2008). Such theories argue that mental representations of language are grounded in perceptual experiences and actions, and that mental simulation (i.e., the mental re-enactment of perceptual, motor and introspective states acquired during experience with the world, body and mind; Barsalou, 2009) is an automatic and integral part of language comprehension.

This raises the question of (1) the nature of the perceptual experiences that underlie voice simulation in response to reading direct versus indirect speech statements in text, and (2) the nature of the representations that are activated during voice simulation.

Regarding the first question, we assume that accumulated experience with direct versus indirect speech *usage* form the basis for voice simulation during silent reading. As discussed in the introduction, when speakers employ direct speech, they often mimic or dramatise aspects of the reported speaker's voice in order to demonstrate or depict the reported speech act; indirect speech, by contrast, is typically not used in such a vivid, demonstrative fashion because its pragmatic function is to provide descriptions of what was said (Clark & Gerrig, 1990). Comprehension of direct speech is therefore more likely to be grounded in the perceptual experience of vivid vocal depictions of the reported speech acts. This would explain why silent reading of direct speech is more likely to engender mental simulations of voice than silent reading of indirect speech.

The second question concerning the exact nature of the simulated voice representations is more difficult to answer at present. We conjecture that the simulated voice representations, as well as their neural correlates, overlap to a large degree with those activated during 'encouraged' auditory imagery (see earlier discussion). However, compared to the speaker-specific voices that are likely to be activated during auditory imagery tasks (particularly following pre-exposure to concrete speech samples), the simulated voice representations reflected in the present study may be less specific, i.e., they may only involve *speaker-unspecific aspects* of voice. The reason for this assumption is that our experimental setup did not encourage the imagination of speaker-specific voices. One of those speaker-unspecific aspects of voice that underlie the present voice simulation findings could be emotional prosody (supra-segmental acoustic information characterising emotional information), which has not only been found to be associated with a right-lateralised activation pattern (cf. Mitchell, Elliott, Barry, Cruttenden, & Woodruff, 2003; Wiethoff et al., 2008; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006), but would also fit well with the notion of direct speech as vivid demonstration (the reported speaker's emotional state is often demonstrated via voice mimicking in direct speech reporting). That said, the exact nature of the voice representations that are spontaneously activated during silent reading of direct versus indirect speech still remains an important question for future research. Part of this has been explored in Chapter 3.

From a broader perspective, it is worth noting that the right STS area (not restricted to the clusters that we observed in the current experiment) is also partially involved in other cognitive processes such as biological motion processing (Sokolov et al., 2011), dynamic face processing (Pitcher, Dilks, Saxe, Triantafyllou, & Kanwisher, 2011), judging the trustworthiness of faces (Dzhelyova, Ellison, & Atkinson, 2011), categorising spectrally complex sounds (Klein & Zatorre, 2011), intention processing in communication (Enrici, Adenzato, Cappa, Bara, & Tettamanti, 2011), perspective taking in Theory of Mind (ToM) tasks (Dodell-Feder, Koster-Hale, Bedny, & Saxe, 2011) and eye gaze processing in social interactions (Hooker et al., 2003). These findings indicate that the right STS area plays an important role in human communication across different sensory and motor modalities.

In conclusion, the present experiment showed that even without being explicitly encouraged to imagine voices, readers are more likely to mentally simulate, or spontaneously imagine, aspects of the reported speaker's voice during silent reading of direct speech as opposed to meaning-equivalent indirect speech. The results can be interpreted in line with embodied cognition (e.g., Barsalou, 1999) and the notion that direct speech is represented in a more vivid and perceptually engaging fashion than indirect speech (Clark & Gerrig, 1990). The present study is novel and important in several respects. Indeed, it is the first demonstration that top-down activation of voice-sensitive areas in the auditory cortex (Belin et al., 2000) can be modulated by pragmatically different reporting styles. Second, it further examines on voice simulations during silent reading. Our results are consistent with the embodied cognition hypothesis, extending it to the auditory perceptual modality which so far has received little attention in the relevant literature (except for studies that focused on sound-related words, see Kellenbach et al., 2001; Kiefer et al., 2008). Third, this study combined event-related fMRI with eye-tracking to investigate neural correlates during on-line reading of text under relatively natural presentation conditions (contrasting with word-by-word reading, for example); this could inspire new applications in psycholinguistic research on language comprehension, helping us to understand the kinds of mental representations activated during reading. Finally, it sheds new light on the distinction between direct versus indirect speech from a cognitive neuroscience perspective, suggesting that perceptual vividness is one of the key aspects differentiating the two.

Chapter 3 Brain 'talks over' boring quotes: Top-down activations of voice-selective areas when listening to monotonous direct speech utterances (Experiment 2)

3.1 Introduction

Combining event-related fMRI and eye-tracking, we found that silent reading of direct speech elicited higher brain activity in voice-selective areas (Belin et al., 2000) of the right auditory cortex than silent reading of indirect speech (Chapter 2). The findings are in line with the demonstration theory of direct speech (Clark & Gerrig, 1990) and the embodied theories of language processing (e.g., Barsalou, 1999; Zwaan, 2004), suggesting that comprehension of direct speech is grounded in the experiences of more vivid vocal depictions and that individuals are more likely to mentally simulate reported speakers' voices during silent reading of direct speech as opposed to indirect speech.

However, with no acoustic stimulation as a reference, it is still unclear what constitutes such 'inner voice' experiences during *silent reading* of direct as opposed to indirect speech. In Chapter 2, we speculated that the mentally simulated voice representations entail supra-segmental acoustic information of the quoted speaker's voice (e.g., speech melodies, intonation and emotional prosody), given that a right-lateralised activation pattern was observed. Indeed, 'bottom-up' auditory stimulation studies have shown the same lateralisation in comparisons of (1) speech or music with acoustically matched noise bursts (Zatorre, Evans, & Meyer, 1994; Zatorre, Evans, Meyer, & Gjedde, 1992), (2) speech signals (irrespective of intelligibility) with noise-vocoded signals (Scott, Blank, Rosen, & Wise, 2000), and (3) nonverbal sounds comprising extended frequency transitions (supra-segmental) with those comprising rapid frequency transitions (sub-segmental) (Johnsrude, Penhune, & Zatorre, 2000). Hence, it seems likely that the right superior temporal gyrus/sulcus (STG/STS) areas are involved in processing dynamic pitch variations which are an important property of supra-segmental vocal information. One type of such information, namely emotional prosody and intonation, is also found to activate similar, right-lateralised activation patterns in various forms including sentences (Mitchell, Elliott, Barry, Cruttenden, & Woodruff, 2003; Wildgruber et al., 2005), words (Wiethoff et al., 2008) and word-like vocalisations (Grandjean et al., 2005). Most importantly, mental simulations of supra-segmental acoustic information in language comprehension would fit

well with the notion of direct speech as vivid demonstration - in which vivid depictions of the quoted speaker's voice are characterised in terms of enriched supra-segmental acoustic information.

In this chapter, we attempted to address whether the mentally simulated voices predominantly consist of supra-segmental acoustic information during comprehension of direct as opposed to indirect speech. In order to create acoustic references to verify such supra-segmental information, we prepared audio recordings of short stories in which direct and indirect speech utterances were spoken monotonously. This manipulation preserved (sub)-segmental acoustic information such as the phonological information of the uttered words in the recordings, but minimised supra-segmental acoustic information such as the global intonation patterns over the utterances. Thus, if direct speech is represented in enriched supra-segmental acoustic representations of voices during language comprehension, individuals would have to mentally simulate such representations, since they are minimised in the stimuli, to supplement what they hear while listening to monotonously spoken direct speech utterances. By contrast, indirect speech is not represented in vivid voice representations and hence individuals need not simulate supra-segmental acoustic information when listening to monotonous indirect speech utterances. Thus, we predict that listening to monotonous direct speech quotations would elicit higher 'top-down' brain activity in voice-selective areas of the right auditory cortex (i.e., similar to the brain areas identified in Chapter 2) than listening to equally monotonous, meaning-equivalent indirect speech quotations.

To test this hypothesis, we employed functional magnetic resonance imaging (fMRI) to measure participants' brain activity while they were listening to short stories which contained monotonously spoken direct or indirect speech utterances. Between the two conditions, we compared the evoked BOLD signal changes within participants' voice-selective brain areas. Moreover, we performed multiple parametric modulation analyses to verify the underlying source of any differential brain activity we observed. To assess the between-study consistency, we also compared the observed activation patterns with our previous silent reading data (Chapter 2).

3.2 Method

3.2.1 Participants

Twenty-one adult participants were recruited and scanned. They were native English speakers with normal hearing and language abilities, and with no history of neurological or psychiatric disorders. Three participants had to be excluded from analysis due to either (a) no clear response in the voice localiser task (1 subject) and/or excessive head-movements during scanning (2 subjects), or (b) scanning abortion following claustrophobic symptoms (1 subject). Data from the remaining 18 participants (age 18 – 32 years, 9 males and 9 females) were valid for the final analyses. All of them were right-handed except for one female subject. They signed the consent form and were paid at £6/hour for their participation.

3.2.2 Stimuli

Listening stimuli

Ninety short stories with different protagonists (indicated by different names) were recorded as stimuli. The stories were exactly the same as in Chapter 2. The full transcriptions are listed in Appendix 1. Each story started with two declarative sentences to set up a scenario (e.g., *Luke and his friends were watching a movie at the cinema. Luke wasn't particularly keen on romantic comedies, and he was complaining a lot after the film.*), followed by either a direct speech or an indirect speech quotation sentence (e.g., *He said: 'God, that movie was terrible! I've never been so bored in my life.'* or *He said that the movie was terrible and that he had never been so bored in his life.*). The reported clauses in both conditions (underscored in the above examples) were equivalent in terms of linguistic content. Additional comprehension questions were also recorded for 23 stories (ca. 25%) to assess participants' overall comprehension accuracy and to ensure that they listen to the stories for comprehension.

The stories and questions were spoken by a professional actress. Critically, in one condition, the direct speech utterances were deliberately spoken as monotonously (*Direct-monotonous* condition) as the indirect speech utterances (*Indirect* condition), that is, without providing vivid depictions of the reported speaker's voice. We also recorded

'normal' (i.e., vivid) versions of the direct speech utterances which were used as a control condition (*Direct-vivid* condition). Example recordings are available at: [www.psy.gla.ac.uk/~boy/fMRI/samplerecordings/].

Three lists of stimuli with counterbalanced item-condition combinations (i.e., 30 Direct-monotonous trials, 30 Indirect trials, and 30 Direct-vivid trials per list) were constructed using a Latin square. Each item appeared once per list, but in a different condition across lists. Each list was assigned to one third of our participants. The presentation order of the items per list was randomised for each participant.

Voice localiser stimuli

For the voice localiser session (see 3.2.3 Procedure), we presented blocks of vocal sounds and non-vocal sounds provided by the Voice Neurocognition Laboratory (vnl.psy.gla.ac.uk), University of Glasgow. These stimuli were the same as those employed in Belin et al. (2000), and comprised both speech (e.g., spoken vowels) and non-speech (e.g., laughing and coughing) vocal sound clips, as well as non-vocal sound clips (e.g., telephone ringing and dog barking). The contrast in brain activity elicited by vocal versus non-vocal sounds reliably localises voice-selective areas of the auditory cortex.

3.2.3 Procedure

Participants were positioned in the scanner, wearing MRI-compatible, electrostatic headphones (NordicNeuroLab, Norway) for (1) auditory presentation during both the story listening session and the voice localiser session and (2) noise attenuation during fMRI scanning. For the story listening session, participants were instructed to keep their eyes closed, to listen to the stories carefully and to answer comprehension questions which would follow 25% of the short stories they had heard. The stimuli were presented using E-prime 2.0 (Psychology Software Tools, Inc., USA); each trial started with a 4-second silence period, followed by the presentation of the story and then (on 25% of the trials) a comprehension question regarding the content of the preceding story. Each such question appeared one second after termination of the preceding story presentation and prompted a 'yes' or 'no' response which participants could provide by pressing buttons on a response box with their index or middle fingers, respectively. The 90 listening trials were evenly

interspersed with five 30-second 'baseline' trials during which no experimental stimulation was present.

After the story listening session, an anatomical scan of the participant's brain was performed, followed by a brief (ca. 10-min) voice localiser scanning session. During the latter, participants were instructed to close their eyes while listening to 20 8-sec blocks of vocal and 20 8-sec blocks of non-vocal auditory stimuli presented in an efficiency optimised, pseudo random order along with 20 8-sec blocks without stimulation, acting as a baseline (cf. Belin et al., 2000).

3.2.4 MRI acquisition

Scanning was performed on a 3-T Siemens Tim Trio MRI scanner using a 12-channel head coil (Erlangen, Germany). Functional scans (for both the story listening session and voice localiser session) were acquired using a T2*-weighted echoplanar imaging (EPI) sequence (32 slices acquired in orientation of the Sylvian fissure; TR = 2s; TE = 30 ms; matrix size: 70 × 70; voxel size: 3 × 3 × 3 mm; FOV = 210). T1 whole-brain anatomical scans were obtained using 3D T1-weighted magnetization prepared rapid acquisition gradient echo (MP-RAGE) sequence (192 axial slices; matrix size: 256 × 256; voxel size: 1 × 1 × 1 mm; FOV = 256). The average scanning time for the whole experiment was around 55 min per participant.

3.2.5 Data analysis

3.2.5.1 Whole brain and ROI analyses

All MRI data were analysed using SPM8 (www.fil.ion.ucl.ac.uk/spm/, University College London). Pre-processing of functional scans included (a) head motion corrections (tri-linear interpolation) whereby scans were realigned to the first volume; (b) co-registration of functional scans to their corresponding individual anatomical scans; (c) segmentation of the co-registered scans; (d) normalisation of functional (3 mm isotropic voxels) and anatomical (1 mm isotropic voxels) data to Montreal Neurological Institute (MNI) space; (e) smoothing of normalised data (10-mm Gaussian kernel).

fMRI data from the anatomical and voice localiser scanning sessions were used to determine the voice-selective areas in the auditory cortex (ROI). The individual voice localisers for most participants (12 subjects) were obtained at $p < 0.05$ (FWE-corrected). The voice localisers for the other 6 subjects were obtained at $p < 0.001$ (uncorrelated, to increase sensitivity). The group voice localiser was obtained at $p < 0.05$ (FWE-corrected).

For the story listening session, the temporal onset of a critical fMRI event was defined as the temporal onset of the first word within the quotation marks (direct speech) or the complementiser *that* (indirect speech); its offset was defined as the temporal offset of the last word in the direct or indirect speech quotations. Non-critical events (e.g., listening to background sentences, comprehension questions and instructions, as well as button pressing) were specified as events of no interest in the design matrix. The rest consisted of all 'silence' events (including five 30-second baseline trials and all 4-second pre-trial silence periods) and were regarded as baseline. The fMRI data were mapped to the human Colin atlas surface (D. C. Van Essen, 2002) in CARET (D. C. Van Essen et al., 2001). The mean beta estimates within ROIs were calculated by SPM toolbox easyROI (www.sbirc.ed.ac.uk/cyril/cp_download.html), and submitted to 2-tailed paired-samples *t*-tests.

3.2.5.2 Parametric modulation analyses

To verify the underlying source of the observed brain activations, we performed parametric modulation analyses with (1) the acoustic parameters, (2) the perceived vividness and (3) the perceived contextual congruency of the critical direct and indirect speech utterances. The acoustic parameters were intended to *objectively* capture the acoustic characteristics of the critical audio recordings independent of linguistic content. The parametric modulations with these measures would unveil whether the observed brain activations were simply engendered by the acoustic differences between conditions. Comparably, the vividness ratings (see below) were intended to provide a more *subjective* measurement of the acoustic characteristics of the critical direct and indirect speech utterances independent of linguistic context. In a way, this measure summarises the joint effect of the acoustic characteristics (i.e., vocal features) and linguistic content (wording etc.) on how vivid the critical utterances would sound to perceivers irrespective of context. Thus, the parametric modulations with the vividness measure would reveal whether between-condition

differences reflect evoked brain responses to the differential vocal features that are subjectively perceived 'bottom-up' from the stimuli. Finally, the contextual congruency ratings were intended to measure the discrepancy between the actual vocal features of the critical stimuli and the way these stimuli 'should have sounded like' in the given contexts. In other words, instead of quantifying the 'bottom-up' perceived vocal vividness of the stimuli, the contextual congruency metric was intended to capture the contextually expected vividness (or its discrepancy with the actually perceived vividness) of the critical speech stimuli. In this sense, parametric modulation analyses with the contextual congruency metric would indicate whether observed brain activation patterns reflect 'top-down' mental simulations of enriched vocal depictions while listening to the critical (monotonous) direct and indirect speech utterances.

Acoustic parameters

Using Praat software (Boersma & Weenink, 2010), we characterised the critical audio samples in terms of eight acoustic parameters which are known to be related to speech prosody: (1) the *mean pitch* (fundamental frequency F_0 averaged over the duration of each sample; (2) the *pitch variation*, measured as the standard deviation in pitch over the duration of each sample (pitch *SD*); (3) the *pitch range* (difference between maximum and minimum pitch in Hz over the duration of each sample); (4) the *mean intensity* (in dB) over the duration of each sample; (5) the *intensity variation*, measured as the standard deviation in intensity over the duration of each sample (intensity *SD*); (6) the *intensity range*; (7) the duration of the voiced sample (the recording periods in which the pitch value passed the voicing threshold); (8) the duration of the entire audio sample. These eight parameters were then included simultaneously as modulators in the parametric modulation analyses to partial out their joint contribution to the between-condition differences in the evoked BOLD signals[6].

[6] Since we are solely concerned with the acoustic parameters' *joint contribution* to the brain activation patterns of interest (and not with each parameter's individual importance), multicollinearity is not an issue here.

Vividness ratings

We recruited 12 native speakers of English with normal hearing and language abilities for this rating study. They were paid at £2 for their participation. A typical session took 10 – 20 min.

Participants were seated in front of a Dell Duo Core PC, wearing headphones. They were presented with the same auditory stimuli that were used in fMRI scanning; only the critical direct and indirect speech utterances (i.e., without context) were presented. After hearing each utterance, participants had to give a rating by pressing number keys on a keyboard, to indicate how vivid and engaging the utterance they had just heard was. The ratings were given in a 7-point scale in which 7 meant 'very vivid and engaging' while 1 meant 'extremely monotonous'.

The collected ratings were summarised by condition for each subject and then submitted to paired-sample *t*-tests to assess the between-condition differences in vividness. They were also summarised by trial and were then included as a parametric modulator to partial out the contribution of perceived vividness to the between-condition differences in the evoked BOLD signals.

Contextual congruency ratings

We recruited a different set of 12 native speakers of English with normal hearing and language abilities for this rating study. They were paid at £4 for their participation. A typical session took 30 – 40 min.

The procedure of this rating study was the same as in the above vividness rating study except: (1) participants were presented with the whole stories (i.e., with context), and (2) they had to give 7 point-scale ratings on the 'contextual congruency' of the critical direct and indirect speech utterances, i.e., on whether these utterances matched the context in terms of how vivid and engaging they were - 7 meant 'fits in the context extremely well' while 1 meant 'does not fit in the context at all'.

The collected ratings were first summarised by condition for each subject and were then submitted to paired-sample *t*-tests to assess the between-condition differences in contextual

congruency. They were also summarised by trial and were then included as a parametric modulator to partial out the contribution of contextual congruency to the between-condition differences in the evoked BOLD signals.

Parametric modulations

We performed three parametric modulation analyses with the speech utterances' acoustics, vividness and contextual congruency as the modulators, respectively. This was to assess each set of modulators' individual contributions to the observed brain-activation differences between the *Direct-monotonous* condition and the *Indirect* condition. Our strategy was to examine whether the differential brain activations would be *reduced* (i.e., accounted for) as a consequence of partialling out the effects of the investigated modulators. First, we performed parametric modulations at the individual level. For each participant, we specified in the design matrix a single trial-type for all three conditions (including the control condition Direct-vivid); it was first followed by one of the three sets of modulator(s) – this would ensure that the effects of the investigated modulator(s) are partialled out across all trials – which were then followed by three experimental conditions coded with binary values. After the experimental trials, other event types and participants' head motion parameters were also included in the design matrix for modelling the collected BOLD signals. We then conducted the contrast analyses in the same way as before. Using the same threshold (i.e., $p < 0.05$, FWE-corrected), we examined how the observed brain-activation differences between the *Direct-monotonous* condition and the *Indirect* condition were affected when the effects of each set of modulators were partialled out, respectively.

3.2.5.3 Comparing brain activation patterns with Chapter 2

Moreover, we also compared the brain activation patterns observed in the current study (in the contrast of Direct-monotonous against Indirect conditions) with the previous silent reading study (Chapter 2). The comparison was descriptive: The activation patterns were described using the 3D coordinates (in relation to the standard brain space from the Montreal Neurological Institute) of the peak voxel within each activation 'blob'. We paired the activation 'blobs' with their counterparts between the two studies and compared the peak voxels' coordinates within each pair.

3.3 Results and discussion

3.3.1 Whole brain and ROI analyses

The voice localiser

Consistent with the findings of Belin et al. (2000), we found that the vocal sounds elicited significantly (*t*s > 7.6, *p*s < .02, FWE-corrected) greater activity than non-vocal sounds bilaterally in the STG/STS areas (Figure 3A). The maximum of voice-sensitive activation was located along the upper bank of the central part of the right STS.

The main contrast

Within the right voice-selective area, we found that listening to monotonously spoken direct speech utterances elicited greater BOLD signals than listening to monotonous indirect speech utterances (see Figure 3B). The differential brain activity between the Direct-monotonous and the Indirect conditions were located in voice-selective areas along the posterior, middle and anterior parts of the right STS brain areas. As a whole, no region showed an opposite pattern of activity, and the Direct-monotonous condition was always associated with a greater BOLD signal than the Indirect condition. Both conditions were active against baseline. The between-condition difference was significant: For individual Regions of Interest (ROIs), 2-tailed paired-sample $t(17) = 5.650$, $p < 0.001$; for the group ROI, $t(17) = 4.979$, $p < 0.001$. The two monotonous conditions elicited significantly less BOLD signal in the right voice-selective area than the Direct-vivid condition (Figure 3C).

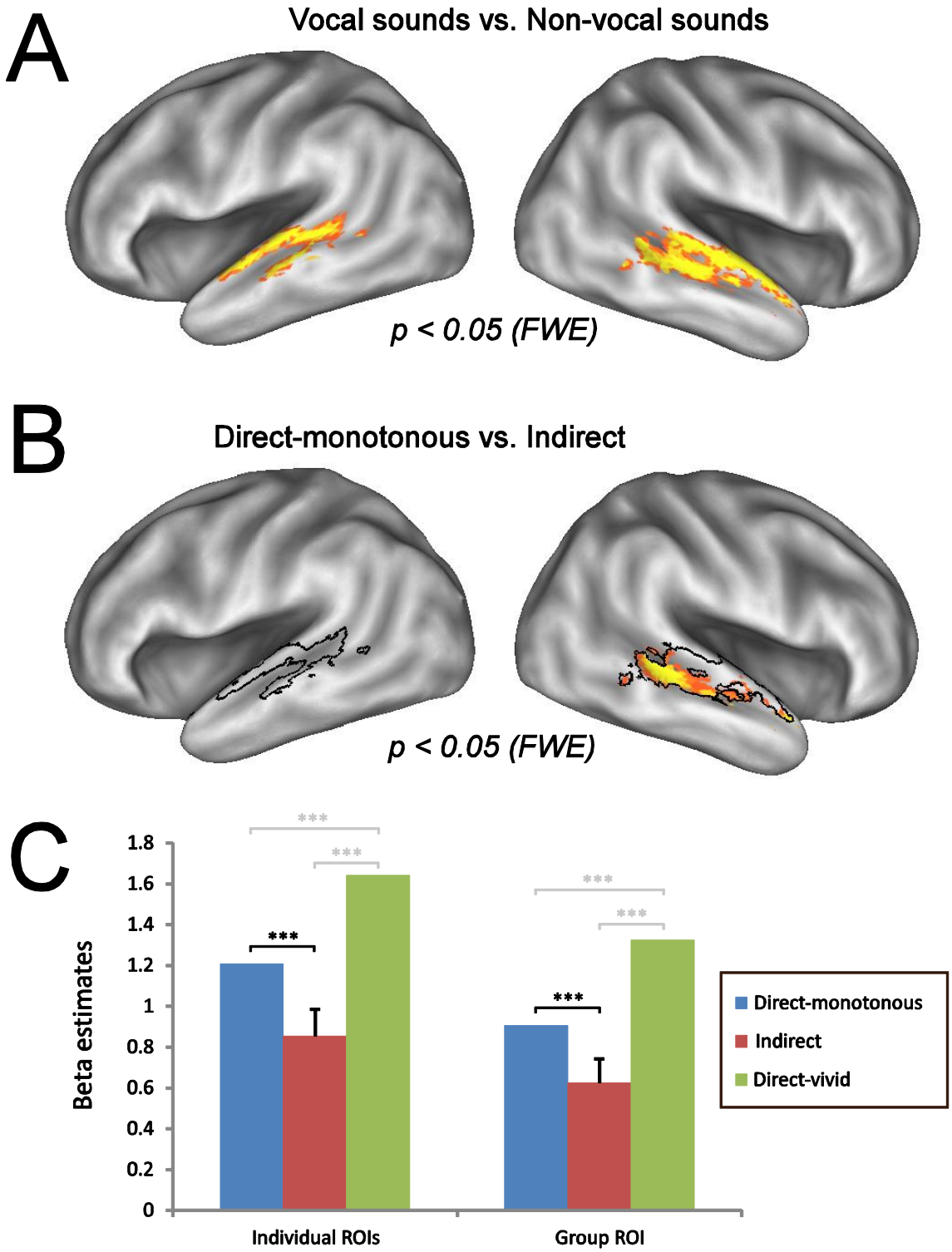


Figure 3 Illustrations of the localizer and the between-condition differences (Direct-monotonous against Indirect), A, Brain regions that selectively responded to vocal sounds as opposed to non-vocal sounds (i.e., the voice-selective areas), **B,** Within the voice-selective areas (indicated by black lines), brain regions that selectively responded to auditory comprehension of monotonous direct speech utterances as opposed to monotonous indirect speech utterances in the whole brain analysis (n = 18), **C,** Mean signal change (against baseline) between conditions in the right voice area Region of Interest (ROIs), determined individually (left) or using the sample average (right). The single error bar in each panel refers to the 95% CI for the between-condition difference (Direct-monotonous against Indirect only).

The higher brain activity in the Direct-vivid condition than in the two monotonous conditions can be explained by the more intense and varied acoustic information that is carried in the more vivid speech utterances (also suggested in Table 2 below). The higher brain activity in the Direct-monotonous condition than in the Indirect condition could be explained in two very different ways. First, it is possible that the Direct-monotonous condition was somehow acoustically more varied than the Indirect condition. In this case, the higher BOLD signals for the Direct-monotonous condition would simply be a reflection of the more varied acoustic information that is carried in those stimuli. However, it is also possible that the Direct-monotonous condition was acoustically *equivalent* to (i.e., not more varied than) the Indirect condition. In that case, the higher BOLD signals for the Direct-monotonous condition must have been due to something other than 'bottom-up' acoustic characteristics of the stimuli, making an explanation more likely whereby monotonous direct speech utterances were supplemented 'top-down' (via mental simulations) with enriched supra-segmental acoustic information.

To distinguish between these contrasting interpretations, we conducted parametric modulation analyses to verify the underlying source of the observed brain activations between Direct-monotonous and Indirect conditions. We investigated the signal contribution of the speech utterances' acoustic parameters, their vividness and their contextual congruency, respectively.

3.3.2 Parametric modulation analyses

It was found that the examined parametric modulators, i.e., the acoustic parameters, the perceived vividness and the perceived contextual congruency, were highly correlated with one another (Table 1). This is not surprising since they all measured variables associated with the vocal vividness of the direct and indirect speech utterances, either objectively, subjectively, or subjectively under consideration of linguistic context (see 3.2.5.2). However, the correlations were certainly not perfect, and the parametric modulation results below revealed rather distinct parametric contributions to the brain activation patterns of interest.

Pearson Correlation (N = 270)	Pitch Mean	Pitch SD	Pitch Range	Loudness Mean	Loudness SD	Loudness Range	Voiced Sample Duration	Recording Duration	Vividness Rating	Contextual Congruency
Pitch Mean	1	.748	.745	.715	.479	.592	.156	.162	.749	.495
Pitch SD	.748	1	.912	.404	.398	.431	.096	.172	.713	.521
Pitch Range	.745	.912	1	.402	.359	.385	.201	.270	.693	.509
Loudness Mean	.715	.404	.402	1	.528	.782	.267	.161	.600	.297
Loudness SD	.479	.398	.359	.528	1	.699	.031	.138	.432	.261
Loudness Range	.592	.431	.385	.782	.699	1	.270	.299	.579	.254
Voiced Sample Duration	.156	.096	.201	.267	.031	.270	1	.856	.232	.135
Recording Duration	.162	.172	.270	.161	.138	.299	.856	1	.288	.224
Vividness Rating	.749	.713	.693	.600	.432	.579	.232	.288	1	.632
Contextual Congruency Rating	.495	.521	.509	.297	.261	.254	.135	.224	.632	1

Table 1 The cross-correlations between the examined parametric modulators.

3.3.2.1 The effects of acoustics

The descriptives of the eight acoustic parameters are summarised by condition in Table 2. It was found that some of the acoustic measures predict *decreased* STS activations in the contrast of the Direct-monotonous condition to the Indirect condition. For instance, the Direct-monotonous condition was associated with lower Pitch SD (therefore less varied acoustic information) as compared with the Indirect condition, and should have elicited *less* brain activity in the voice-selective areas. Therefore, these acoustic measures could not explain the *enhanced* STS activations that we observed in the Direct-monotonous condition as opposed to the Indirect condition. However, acoustic measures related to intensity (i.e., Intensity Mean, Intensity SD and Intensity Range) were found to display a congruent pattern with the observed brain activation difference between the two conditions: The Direct-monotonous condition was associated with significant higher Intensity Mean, SD and Range, paired-sample *t*s(89) > 2.5, *p*s < .013. Although the between-condition differences in these acoustic measures were statistically significant, the sizes of the differences do not seem to be evidently meaningful in practice. For instance, the Intensity Mean difference between the Direct-monotonous and the Indirect conditions amounted to a mere 1.12 dB, which is hardly noticeable in loudness perception, given that a 5 dB increment step size is usually recommended for subjects with normal-hearing sensitivity (Cox, Alexander, Taylor, & Gray, 1997). Nonetheless, if these intensity differences did contribute to the brain-activation differences between the two critical conditions, we expect that by partialling out the effects of these acoustic measures in the following parametric

modulation analysis, the between-condition differences in brain activations should be subsequently accounted for (i.e., reduced).

Acoustic parameters	Condition		
	Direct-monotonous	**Indirect**	**Direct-vivid**
Pitch Mean (Hz)	200.43 (12.42)	208.50 (13.39)	265.81 (43.22)
Pitch SD (Hz)	32.62 (6.97)	45.24 (10.51)	66.36 (20.11)
Pitch Range (Hz)	143.87 (39.27)	199.97 (52.80)	281.77 (82.48)
Intensity Mean (dB)	68.01 (1.17)	66.89 (.93)	70.71 (2.48)
Intensity SD (dB)	7 (.51)	6.87 (.44)	7.51 (.68)
Intensity Range (dB)	31.23 (1.62)	29.66 (1.59)	34.32 (3)
Voiced Duration (ms)	2044 (534)	2073 (559)	2318 (662)
Audio Sample Duration (ms)	4028 (984)	4332 (1056)	4702 (1119)

Table 2 A summary of the eight acoustic parameters (means, with standard deviations in parentheses) for the spoken stimuli in each experimental condition.

The parametric modulation analyses showed that, using the same threshold ($p < 0.05$, FWE-corrected), the Direct-monotonous condition still elicited higher brain activity within the right voice-selective area than the Indirect condition after the effects of the acoustic parameters were partialled out (Figure 6B). Since the activation patterns were highly similar to the original contrast (Figure 6A), it suggests that the observed brain-activation differences between the Direct-monotonous and the Indirect conditions were unlikely to be engendered by the between-condition differences in acoustic characteristics.

3.3.2.2 Partialling out the effect of vividness

The averaged ratings and the paired-sample *t*-test results are illustrated in Figure 4. It was found that the control condition Direct-vivid was perceived as significantly more vivid than the two monotonous conditions, $ts(11) > 19$, $ps < .001$. However, instead of being 'equally monotonous', the Direct-monotonous condition was perceived as significantly *less* vivid than the Indirect condition, paired-sample $t(11) = -7.872$, $p < .001$. The latter suggests that the monotonous direct speech utterances (Direct-monotonous) contained less vivid vocal modulations than the monotonous indirect speech utterances (Indirect). This is also inconsistent with a more 'bottom-up' explanation of our findings (Figure 3): Given lower vividness ratings for the Direct-monotonous condition, a 'bottom-up' account would predict that this condition would consume less energy (therefore less blood oxygen) to process within the voice-selective areas than the more vivid Indirect condition; in other words, the Direct-monotonous condition should have elicited significantly *decreased*

BOLD signals within these brain areas compared to the Indirect condition. However, exactly the opposite was found (see 3.3.1).

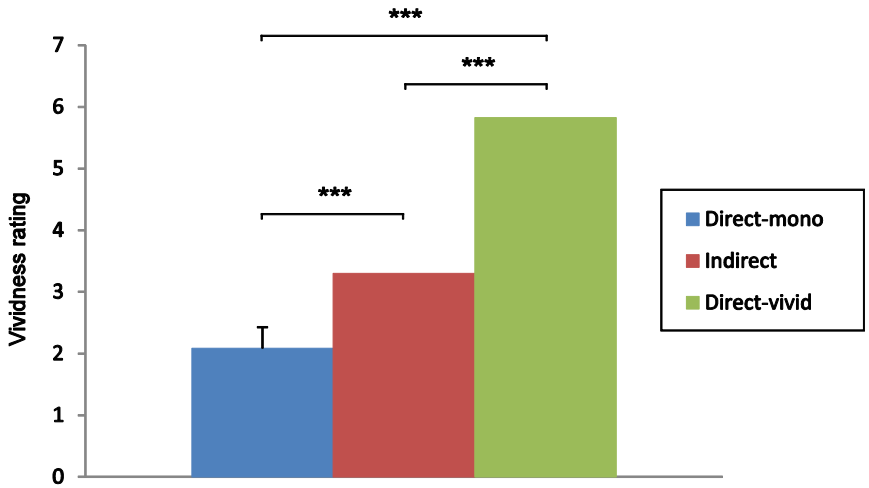


Figure 4 Illustrations of the between-conditions differences in vividness. The significance of the pairwise contrasts (referred to with square brackets) were indicated with asterisks (*** indicates $p < .001$). The single error bar represents the 95% CI of the between-condition difference between Direct-monotonous and Indirect conditions.

The parametric modulation analysis on vividness revealed that using the same threshold ($p < 0.05$, FWE-corrected), the Direct-monotonous condition still elicited higher brain activity within the right voice-selective area than the Indirect condition after the effects of vividness were partialled out (Figure 6C); the between-condition difference appeared to be *larger* and *more widely spread* than the original contrast (Figure 6A). This suggests that the originally observed brain-activation difference between the Direct-monotonous and the Indirect conditions is not 'explained' by the vividness of the speech utterances (for such a conclusion to be justified, there should have been a *reduction* in the original difference after partialling out the effect of vividness). Instead, it indicates that the original between-condition difference was partially 'masked' by the vividness contrast between the two conditions: The vividness ratings indicate that the Direct-monotonous condition was perceived as *less* vivid than the Indirect condition (see Figure 3); when the two conditions were brought to the same vividness level (by partialling out the effect of the vividness modulator), the original brain-activation difference was enhanced. In other words, the influence of factors other than vividness became more pronounced when the negative contribution of vividness (Direct-monotonous < Indirect) was eliminated. We conclude that while perceived vividness clearly played a role in the originally reported brain

activation patterns, its contributions actually went *contrary to* an actual explanation of those brain activation patterns.

3.3.2.3 Partialling out the effects of contextual congruency

The averaged ratings and the paired-sample *t*-test results are illustrated in Figure 5. It was found that while the Indirect and the Direct-vivid conditions were perceived as equally congruent with the preceding linguistic contexts, paired-sample $t(11) = 1.754$, $p > .1$, the Direct-monotonous condition was perceived as significantly less congruent with contexts than both of them, paired-sample $ts(11) > 4$, $ps < .003$ (Figure 5). The results show that listeners routinely expect vivid vocal depictions for direct speech but not for indirect speech; they have to mentally simulate or imagine vivid depictions of the reported speaker's voice to supplement the monotonously spoken direct speech utterances (contextually incongruent) but not the monotonous indirect speech utterances or the vivid direct speech utterances (contextually congruent). This voice simulation process in the Direct-monotonous condition would have required additional energy consumption within the voice-selective areas of the auditory cortex; this could explain why the less vivid Direct-monotonous condition (see 3.3.2.2), which should have elicited decreased BOLD signals in the voice-selective areas of the auditory cortex under a more 'bottom-up' interpretation, actually elicited significantly higher BOLD signals within these brain areas as compared to the Indirect condition.

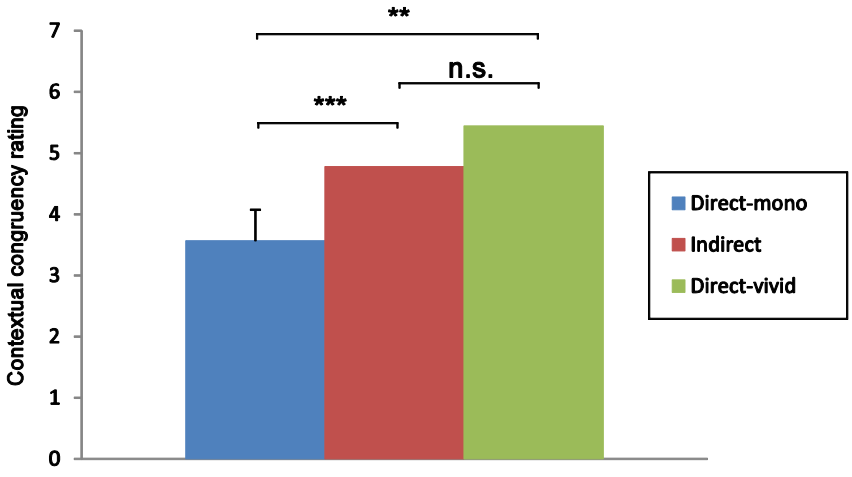


Figure 5 Illustrations of the between-conditions differences in contextual congruency. The significance of the pairwise contrasts (referred to with square brackets) were indicated with asterisks and abbreviations (*** indicates $p < .001$, ** = $p < .01$, n.s. = not significant). The single error bar represents the 95% CI of the between-condition difference between Direct-monotonous and Indirect conditions.

The parametric modulation analyses revealed that under the same threshold ($p < 0.05$, FWE-corrected), only a few voxels at the right pSTS area survived when contrasting the Direct-monotonous condition to the Indirect condition after the effects of contextual congruency were partialled out (Figure 6D); compared to the original contrast (Figure 6A), the between-condition difference had almost disappeared. The results confirm that the observed between-condition difference reflects mental simulations of enriched supra-segmental vocal depictions specifically when listening to monotonously spoken direct speech utterances rather than to monotonous indirect speech utterances.

3.3.2.4 Summary

The parametric modulation results showed that the increased brain activity during auditory language comprehension of monotonous direct speech as opposed to monotonous indirect speech can in part be explained by the contextual congruency of the direct or indirect speech utterances, but not by their acoustic characteristics or their perceived vividness out of context. It suggests that listeners routinely expect vivid depictions of the reported speaker's voice for direct speech but not for indirect speech, and that they are more likely to mentally simulate such enriched supra-segmental vocal representations while listening to direct speech utterances which are spoken monotonously as opposed to monotonous, meaning-equivalent indirect speech utterances.

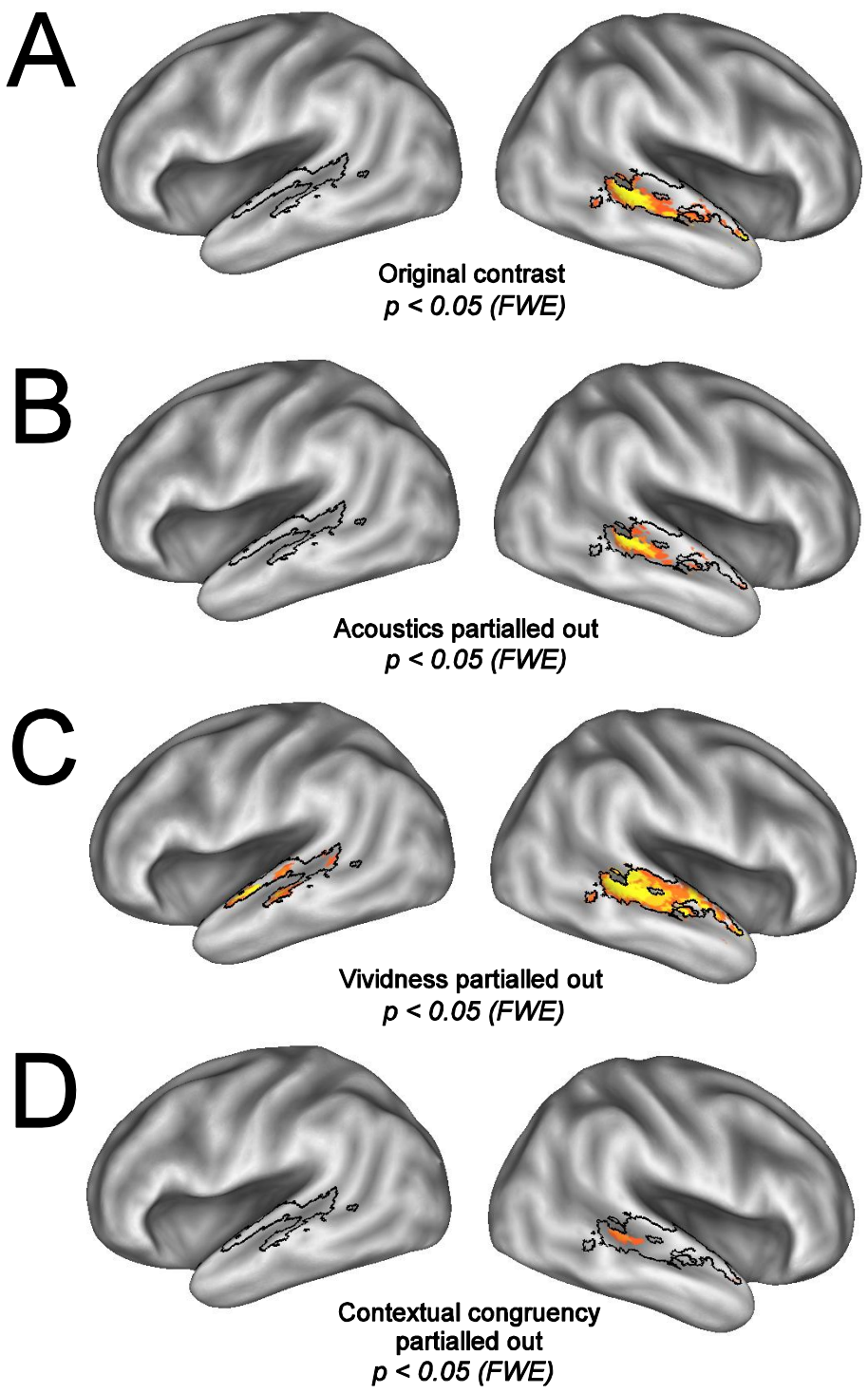


Figure 6 Illustrations of between-condition differences (Direct-monotonous against Indirect) in different parametric modulation models (the voice-selective areas are indicated by black lines), A, the original contrast, **B,** the same contrast with effects of acoustics partialled out, **C,** the same contrast with effects of vividness partialled out, **D,** the same contrast with effects of contextual congruency partialled out.

3.3.3 Comparing brain activation patterns between studies

It was found that the activation patterns observed when listening to monotonous direct speech against monotonous indirect speech resembled those observed in silent reading of direct speech against indirect speech (cf. Chapter 2). The brain activations patterns in both studies were located at the posterior, the middle and the anterior parts of the right STS areas. Within the MNI space, the peak voxels within each activation cluster were spatially close to their counterparts across the two studies (Figure 7). The between-study consistency in the activation patterns suggests that the 'inner voice' we observed in silent reading of direct as well as indirect speech (Chapter 2) is similar in nature to the enrichment of monotonous direct speech (as opposed to indirect speech) that we found in the current study. Given that the 'vocal enrichments' we observed in the present study entailed supra-segmental acoustic information that was hardly available in the actual stimuli, the 'inner voice' we observed in silent reading may also have been supra-segmental in nature.

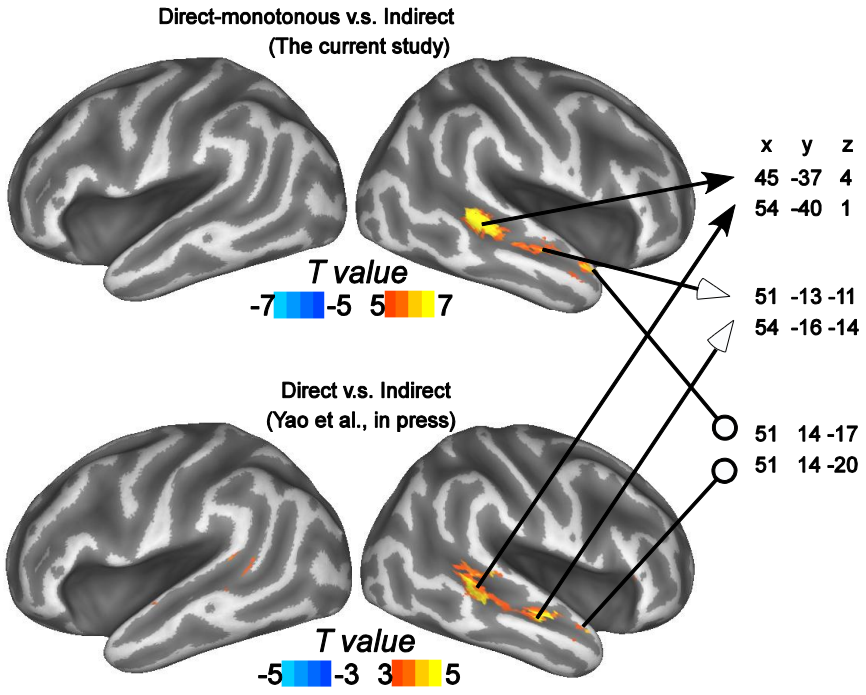


Figure 7 Illustrations of the critical contrasts between the two studies (masked by voice-selective areas). The top panel shows the contrast between Direct-monotonous and Indirect conditions in the current study (listening). **The bottom panel** shows the contrast between Direct speech and Indirect speech in Chapter 2 (silent reading). The arrows point to the peak voxels' 3D coordinates (in MNI space) in the activation clusters. The peak voxels were paired with their anatomical counterparts between the two studies. The thresholds for the two contrasts were adjusted to better illustrate the activation blobs.

3.4 Summary and general discussion

The current experiment investigated mental simulations of supra-segmental acoustic representations during auditory language comprehension of direct as opposed to indirect speech. We employed audio recordings in which direct and indirect speech utterances were spoken monotonously. This manipulation preserved the (sub)-segmental acoustic information (e.g., phonological information associated with individual words) but suppressed the supra-segmental acoustic information (e.g., the intonation patterns across the speech utterances). Using event-related fMRI, we found that listening to monotonously spoken direct speech utterances elicited significantly higher brain activity within voice-selective areas of the right auditory cortex as compared to listening to monotonous, meaning-equivalent indirect speech utterances. Part of the between-condition difference was significantly accounted for by the perceived contextual congruency of the direct and indirect speech utterances but not by the utterances' acoustic characteristics or out-of-context vividness (see 3.2.5.2 for the rationale behind these three measures).

The results might be interpreted in terms of violations of listeners' expectation of vivid vocal depictions for direct speech, which may not necessarily require mental simulations of the expected supra-segmental acoustic information. However, fMRI studies shows that the brain networks associated with violations of expectation typically consist of frontal cortex such as dorsolateral prefrontal cortex (dlPFC) and ventrolateral PFC, lateral parietal cortex and/or the basal ganglia but not the right STS area (Bubic, von Cramon, Jacobsen, Schroger, & Schubotz, 2009; Huettel, Mack, & McCarthy, 2003). Moreover, the detection of discrepancies/incongruencies between the semantic content (e.g., *She won the lottery jackpot*) and the emotional prosody of utterances (e.g., a sad emotional prosody) has been found to activate the left inferior frontal gyrus (Mitchell, 2006), as well as the right dorsal anterior cingulate cortex (dACC) and the right mid STG/STS areas. These studies show that violations of expectation are typically associated with more spreading brain activations distributed among a network of brain areas. Importantly, the frontal brain areas seem to play a critical role in processing conflicts between perception and expectation. Given that in the current experiment, the brain-activation differences were only observed along the right STS areas, the 'expectation violation' account does not fit as well with the existing evidence. One possibility may be that with no explicit incongruency detection task (the current experiment employed a normal language comprehension task), the brain can

resolve the violations of expectation of vivid vocal depictions locally within the auditory cortex without communicating with the frontal cortex. Further research is needed to verify this account.

An alternative and perhaps better interpretation of the results adopts the notion of perceptual simulations (e.g., Barsalou, 1999): Listeners expect vivid vocal depictions (i.e., enriched supra-segmental acoustic information) for direct speech but not for indirect speech; they are more likely to mentally simulate such information (if it is not available in the stimulus itself), thereby effectively supplementing what they hear when listening to monotonously spoken direct speech utterances (contextually incongruent) but not monotonously spoken indirect speech utterances (contextually congruent). Unlike the 'expectation violation' account, this interpretation fits well, both theoretically and empirically, with the previous findings on voice simulations in reading (Yao, Belin, & Scheepers, 2011; Yao & Scheepers, 2011). For this reason, we are leaning towards a 'simulation' account in that the greater right STS activations when listening to monotonous direct speech reflect enhanced mental simulations of vivid vocal depictions (i.e., enriched supra-segmental acoustic information).

The current experiment sheds new light on the nature of the 'inner voice' representations that are mentally simulated in language comprehension of direct as opposed to indirect speech. In the previous silent reading experiment (Chapter 2), there was no experimentally manipulated auditory stimulation which could be used as a reference to the representational nature of the observed 'inner voice' activations. The current study resolved this issue by using monotonously spoken direct versus indirect speech utterances in which the enriched supra-segmental vocal information is scarcely available while the (sub)-segmental acoustic information is intact. This manipulation suggested that the mentally simulated (as opposed to acoustically perceived) voice representations must be supra-segmental rather than (sub)-segmental in nature. Intriguingly, the 'inner voice' activations observed in the current investigation and in Chapter 2 were located in virtually the same brain areas. Reconciling the findings of the two studies, we infer that the 'inner voice' we observed in silent reading of direct as opposed to indirect speech may also entail supra-segmental acoustic representations.

At a broader level, the current findings also provide novel insights into speech perception. Speech perception is not a passive information processing mechanism in that it involves both bottom-up and top-down processes. Various top-down influences in speech perception have been documented: Listeners use prior lexical knowledge or perceptual experience in perceptual grouping of speech, segmenting connected speech, perceptual learning of distorted speech, and perceiving speech categorically (see Davis & Johnsrude, 2007 for a review). However, such top-down influences have mostly been documented in terms of low-level 'speech recognition' - researchers have mainly focused on how listeners use prior lexical knowledge or perceptual experience to interpret *distorted, unintelligible,* or *ambiguous* speech. This neglects the fact that speech perception also involves comprehension of the recognised linguistic information. It is currently less clear whether and how on-line comprehension of auditory linguistic information also influences upcoming *intelligible* and *unambiguous* speech. Moreover, many studies examine the top-down interactivity at the *(sub)-segmental* level (e.g., word interpretation), whereas the top-down influences at the *supra-segmental* level (e.g., emotional prosody) have received limited attention. The current study demonstrated top-down activations of *supra-segmental* acoustic representations during *intelligible* and *unambiguous* speech perception of direct versus indirect speech utterances. It provides evidence that during natural speech perception, top-down interpretations of incoming acoustic signals routinely take place even at the supra-segmental level. More importantly, such top-down interpretations are modulated as a function of linguistically/pragmatically different reporting styles (direct versus indirect speech). Our findings emphasise that in addition to prior lexical knowledge and perceptual experience, other linguistic factors such as reporting style should also be considered in modelling the top-down interactivity of natural speech perception.

In conclusion, the current study shows that listeners routinely expect vivid depictions for direct speech but not for indirect speech; they spontaneously engage in mental simulations of vivid vocal depictions while listening to monotonously spoken direct speech as opposed to monotonously spoken indirect speech. The findings replicate our previous findings of an 'inner voice' during silent reading of direct as opposed to indirect speech, but within a different perceptual modality. This highlights the universality of such voice simulation process in comprehension of direct speech. Furthermore, it provides evidence that the nature of the mentally simulated 'inner voice' entails supra-segmental acoustic

representations. It also verifies the neural correlates of such voice simulation process, which include the anterior, the middle and the posterior parts of the right STS brain areas. Future research should seek to specify the exact function of the brain areas involved in such simulation processes. Finally, from a broader perspective, the current findings extend the scope in modelling natural, intelligible speech perception, emphasising that comprehension-driven, top-down influences at the supra-segmental level should also be considered.

PART 2: The Acoustic Aspects of the 'Inner Voice' in Comprehension of Direct versus Indirect Speech

Using event-related fMRI in both silent reading (Chapter 2) and listening (Chapter 3), we have demonstrated that direct speech elicited higher brain activity in voice-selective areas of the right auditory cortex as compared to indirect speech. The findings suggest that individuals are more likely to mentally simulate enriched supra-segmental acoustic representations of reported speakers' voices during language comprehension of direct speech quotations. However, it was still underspecified what acoustic information constitutes such supra-segmental acoustic representations, and whether there would be potential behavioural consequences of voice simulation during reading (e.g., modulation of eye-movement patterns) and voice perception.

In the second stage of this doctoral research, we employed a range of behavioural methodologies and attempted to address these questions by decomposing the purported 'inner voice' representations into acoustical parameters. Specifically, we focused on three fundamental acoustic parameters of a speaker's voice, i.e., pitch (Chapter 4), loudness (Chapter 5) and rate (Chapter 6), respectively. We examined how individuals' respond to the direct and indirect speech reporting styles as well as the preceding linguistic contexts during both oral and silent reading.

Chapter 4 Contextual modulation of pitch representations in oral and silent reading of direct versus indirect speech

4.1 Introduction

The present chapter examined the pitch dimension of the observed 'inner voice' experience, namely the contextually implied *pitch* of the reported speaker, during language comprehension of direct as opposed to indirect speech. Pitch (i.e., fundamental frequency, F_0) is an important acoustic parameter of a speaker's voice. It carries various types of information about a speaker such as his/her gender (Kent & Read, 1992) and emotional state (Banse & Scherer, 1996; Banziger & Scherer, 2005; Busso, Lee, & Narayanan, 2009). It also signals the structures of speech utterances (e.g., their syntactic and prosodic boundaries; see Beattie, Cutler, & Pearson, 1982; Danly, Cooper, & Shapiro, 1983; Hird & Kirsner, 2002; Menn & Boyce, 1982; Umeda, 1981; Vaissiere, 1983) and what is emphasised in a speech utterance (Pell, 2001). It is also a beneficial cue for decoding speech in noisy environments (Brown & Bacon, 2010). Hence, given the importance of pitch in verbal communication, we hypothesised that the observed 'inner voice' experience in comprehension of direct speech may contain representations of the reported speaker's pitch.

One way of detecting such pitch representations is to examine whether individuals would adjust their own pitch in accordance with the contextually (linguistically) implied pitch of the reported speaker during oral reading of direct versus indirect speech. Oral reading initially requires *comprehension* of the presented written text, which is followed by *oral interpretations* of the formed linguistic representations into strings of vocal sounds (i.e., speech). Since our brain-activation investigations (Chapter 2 & 3) have shown that individuals are more likely to simulate reported speakers' voices during *comprehension* of direct speech rather than indirect speech, we predict that individuals would be more likely to engage in similar voice simulations during *oral reading* of direct speech as opposed to indirect speech; critically, while the simulated voice representations are being orally interpreted into speech, the acoustic characteristics of such representations may be reflected in concordant modulations of vocal acoustics. In other words, if individuals mentally simulate pitch information in accordance with the contextually implied pitch of

the reported speaker during oral reading, they may interpret such information by adjusting their own pitch when reading direct speech, but less so when reading indirect speech.

Experiment 3 was aimed to test this hypothesis. We prepared short stories in which we manipulated the reported speaker's pitch level by manipulating their emotional arousal and gender. On one hand, it has been shown that a speaker's emotional arousal (Banse & Scherer, 1996; Bulut & Narayanan, 2008) and gender (Kent & Read, 1992) are reliably reflected in modulations of pitch (e.g., mean pitch, $\overline{F}_0$) during articulation. For example, $\overline{F}_0$ is positively correlated with the speaker's emotional arousal: High-arousal emotion such as *Anger* or *Joy* is generally associated with increased $\overline{F}_0$ whereas a low-arousal emotion such as *Sadness* is associated with lowered $\overline{F}_0$ (Banse & Scherer, 1996). During oral reading of direct speech rather than indirect speech, if individuals mentally simulate reported speakers' pitch information, they may increase their own pitch when they read about a reported speaker who is in a high-arousal emotion (e.g., *Anger*) as opposed to when they read about a speaker in a low-arousal emotion (e.g., *Sadness*). On the other hand, $\overline{F}_0$ also indicates a speaker's gender: Women tend to speak in a higher pitch than men because of different anatomical structures of their vocal tracts (Kent & Read, 1992). Hence, we also predicted that individuals may increase their pitch during oral reading of direct rather than indirect speech when the reported speaker is female (e.g., *She said*) as compared to when the reported speaker is male (e.g., *He said*).

In addition to the above pitch manipulations, we also manipulated linguistic focus on reported speakers' voices in order to verify the nature of the representations that underlie any observed contextual modulation of $\overline{F}_0$ during oral reading of direct versus indirect speech. Linguistic focus modulates mental representations during language comprehension (Sturt, Sanford, Stewart, & Dawydiak, 2004; Taylor & Zwaan, 2008). For example, Taylor and Zwaan's previous findings (Zwaan & Taylor, 2006) demonstrated that written descriptions of manual rotation (e.g., *Eric turned down the volume*) activate compatible motor representations (i.e., an anticlockwise manual rotation) which interacted with participants' concurrent manual rotation (for self-paced reading, using a knob). They named such effects (i.e., the interactions between reading and manual rotation) *motor resonance*. They found that motor resonance was localised to the verb that denoted the action (i.e., *turned down*) and decayed quickly during reading of the post-verb information.

In their later study (Taylor & Zwaan, 2008), they explored whether mental simulations of actions in reading were maintained while the actions were within linguistic focus and were diminished as linguistic focus shifted away. They found that when a postverbal adverb maintains linguistic focus on a matching action (e.g., 'slowly' in relation to 'turned down the volume'), motor resonance is accordingly maintained (i.e., mental representation of an anticlockwise rotation is maintained); when a postverbal adverb shifts the linguistic focus to the agent (e.g., 'eagerly'), motor resonance ceases. Their results suggest that motor simulations are yield to linguistic focus. Thinking along the same lines, we postulated that voice simulations may also be modulated by linguistic focus. The rationale here is that if the purported voice simulations during oral reading of direct speech (should be reflected in pitch modulations by reported speakers' emotional arousal and gender) can be *selectively* modulated by exerting linguistic focus on voices (as opposed to non-vocal aspects such as faces), we would have evidence that their underlying representations are likely to be voice-related and to contain pitch information. Thus, in Experiment 3, we manipulated linguistic focus on voices and investigated its effects on voice simulations during reading. Because the critical direct speech and indirect speech quotations were always the final sentences in our stimuli, we could not manipulate linguistic focus after quotations (i.e., the critical regions for simulations) as Taylor and Zwaan had. As an alternative, we manipulated linguistic focus before the critical quotations by using voice-focused (e.g., *His voice was furious and raspy: '...'*) versus voice-unfocused (e.g., *He said, with his face turning red: '...'*) linguistic descriptions. Assuming that exerting linguistic focus on reported speakers' voices has an effect on the subsequent oral reading of direct and indirect speech quotations. In principle, it could have two contrasting effects. One possibility is that linguistic focus on voices may enhance mental simulations of voices during oral reading of direct speech but not indirect speech. In other words, the expected emotion effects (i.e., the emotional arousal of the reported speakers) and gender effects (i.e., the gender of the reported speakers) on participants' reading pitch may be more robust when a direct speech quotation is preceded with a voice-focused description as opposed to when it is preceded with a voice-unfocused description. Thus, we should observe that the pitch modulation difference between oral reading of direct speech and indirect speech quotations would be *increased* when the linguistic focus is exerted on reported speakers' voices (see Figure 8A). The other possibility is that the linguistic focus on voices does not necessarily enhance voice simulations in oral reading of direct speech because direct speech quotations alone

can trigger voice simulations sufficiently (a ceiling effect), but it enhances voice simulations in oral reading of indirect speech (see Figure 8B). Thus, the difference in pitch modulations between oral reading of direct speech and indirect speech quotations would therefore be *reduced* or *"masked"* when linguistic focus is exerted on reported speakers' voices. While linguistic focus on voices is expected to modulate voice simulations in oral reading of direct versus indirect speech, linguistic focus on non-vocal aspects (i.e., the direct and indirect speech quotations are preceded with voice-unfocused descriptions) should not affect the pitch modulations in oral reading of direct versus indirect speech. If we observe such *selective* modulations (either enhancing or masking) of voice simulations in oral reading by exerting the linguistic focus on *voices*, we could conclude that the underlying representations of the observed pitch modulations during oral reading of direct versus indirect speech are likely to be voice-related in nature and may contain pitch information.

In sum, Experiment 3 investigated the pitch aspect in mental simulations of voices during oral reading of direct versus indirect speech. We examined whether individuals adjust their own pitch ($\overline{F}_0$) in accordance with the reported speaker's emotional arousal and gender that are contextually (linguistically) implied in the preceding context when reading direct speech but not indirect speech. We also manipulated the linguistic focus on reported speakers' voices to verify the voice-related nature of the underlying representations of any observed $\overline{F}_0$ modulations during oral reading of direct versus indirect speech. Participants orally read written stories which contained direct or indirect speech quotations while audio recordings of their voices were taken for acoustic analyses.

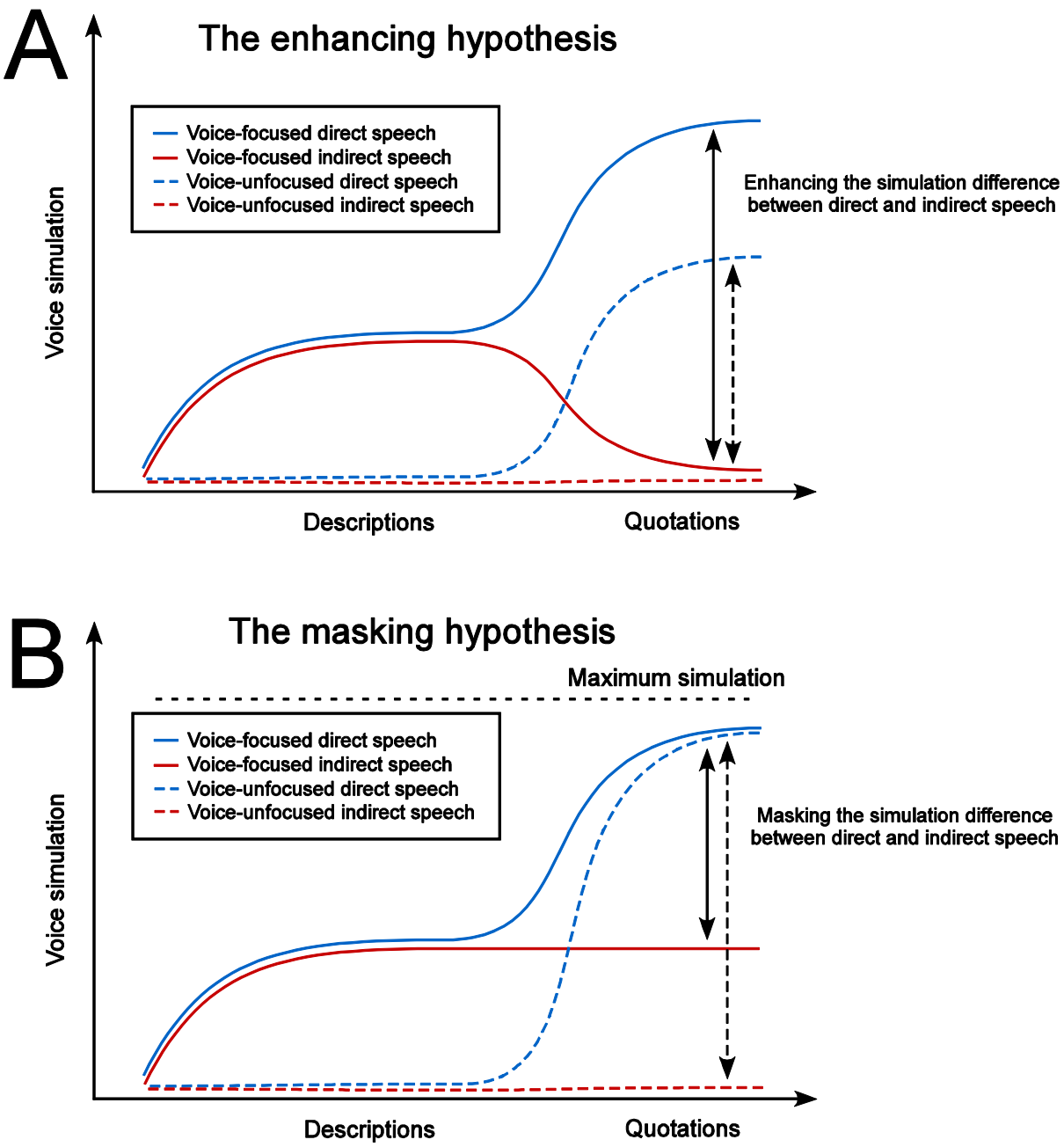


Figure 8 Illustrations of the potential effects of linguistic focus on voices on voice simulations in oral reading of direct versus indirect speech. A, the illustration of the 'enhancing' effect, **B,** the illustration of the 'masking' effect.

The *oral reading* paradigm is useful in exploring acoustic parameters of the 'inner voice' experience by examining individuals' articulation during oral reading of direct versus indirect speech. However, it is also important to explore mental simulations of such acoustic parameters during *silent reading* of direct versus indirect speech (i.e., without the demand for articulation). One way of studying mental representations in silent reading is to adopt the reaction time-based 'reading & recognition' paradigms that have been

commonly used in the literature (e.g., Kurby et al., 2009; Stanfield & Zwaan, 2001; Zwaan et al., 2002). The 'reading & recognition' paradigms typically consist of a main reading task and a secondary recognition task. They examine the effects of language comprehension (in the reading task) on subsequent recognition of the relevant perceptual stimuli (in the recognition task): If the relevant perceptual representations (e.g., a horizontal orientation of a nail) are mentally simulated during reading (e.g., of *He hammered the nail in the wall*), they are expected to differentially influence individuals' recognition latencies (i.e., reaction times) to the perceptually-congruent (e.g., a picture depicting a horizontally presented nail) versus -incongruent probe stimuli (e.g., a picture depicting a vertically presented nail). Using similar paradigms in our investigations, we predicted that if individuals mentally simulate reported speakers' pitch representations (e.g., a high-pitch voice of a woman) during silent reading of direct rather than indirect speech, they have an advantage in recognising pitch-congruent stimuli (e.g., a high-pitch voice of a woman) as opposed to pitch-incongruent stimuli (a low-pitch voice of a man).

We developed three versions (Experiments 4 - 6) of such 'reading & recognition' paradigms to explore the hypothesised mental simulations of pitch during silent reading of direct versus indirect speech. Because that we did observe noticeable interaction between linguistic focus and the *Reporting Style* × *Emotion* interaction in Experiment 3 (i.e., the underlying representations of the *Emotion* effects may not be voice-related), as well as for the purpose of experimental design simplification, we exclusively focused on mental simulations of pitch information in relation to reported speakers' genders in our silent reading investigations. We examined the effects of genders on the recognition latencies to a spoken probe word when its speaker's gender was made relevant by the previous text passage. Participants silently read a series of short stories containing either a direct speech or an indirect speech, each followed by a spoken probe word, and answered 'yes' or 'no' to the spoken word on a secondary task. The probe words were spoken by either a male or a female speaker such that it either matched or mismatched the reported speaker's gender described in the preceding story. Just as actually hearing a voice would facilitate recognition of a congruent voice later, so too should simulating the voice facilitate later recognition of a congruent voice. The greater the congruency between the simulated voice and the heard voice (i.e., the more acoustic features they have in common), the greater the degree of facilitation. If pitch representations are mentally simulated in accordance with

the reported speakers' genders during silent reading of direct as opposed to indirect speech, recognition of probe words should be facilitated when the words are spoken by a 'same-gender' (i.e., same as the reported speaker's gender) speaker as compared to when they are spoken by an 'opposite-gender' speaker because the former are acoustically more congruent with the simulated voices in pitch. Thus, we hypothesised that individuals should respond faster to words that are spoken by a 'same-gender' speaker than to words spoken by an 'opposite-gender' speaker immediately after or during silent reading of direct speech but not indirect speech. Experiment 4 employed a self-paced reading task, followed by a spoken word recognition task in which participants had to indicate whether the spoken word had been mentioned in the story they had just read. Experiment 5 utilised the same self-paced reading task, which was, however, followed by a different lexical decision task in which participants had to indicate whether the spoken word was a meaningful English word or not. The lexical decision task was believed to exert a significantly lighter memory demand as compared to Experiment 4's recognition task[7]; it was intended to retain the natural reading process (which does not require a heavy memory load) as much as possible. While both Experiment 4 and 5 probed the simulated voice representations off-line (i.e., the spoken words were presented 1000 ms after silent reading of text), Experiment 6 employed an serial visual presentation (SVP) paradigm in which the written critical words and the spoken words were presented simultaneously. This paradigm was designed to maximise the sensitivity in detecting voice representations. It has been shown that mental simulations of motor representations during language comprehension are dynamic and short-lived (Borreggine & Kaschak, 2006; Glenberg et al., 2008; Zwaan & Taylor, 2006); hence it is possible that the simulated voice representations are also short-lived and will decay quickly after reading. If this was the case, the off-line paradigms used in Experiments 4 & 5 may not be sensitive enough to detect the simulated voice representations – the latter may have already decayed before the presentation of spoken probe words. By contrast, simultaneous presentation of the written critical words (for reading) and the spoken probe words (for recognition) would be able to promptly detect even the short-lived representations of voices.

[7] The recognition task requires memorisation of every single word of a story. Such a heavy memory load may disrupt normal reading process.

In sum, Experiments 4 to 6 investigated the pitch aspect in voice simulations during silent reading of direct as opposed to indirect speech. We examined whether silent readers would respond differentially to words that were spoken by a same-gender speaker (i.e., the same as the reported speaker's gender) versus by an opposite-gender speaker. Participants silently read short stories and responded to spoken words immediately after or during silent reading of direct or indirect speech quotation in the story. Their response times and accuracy were recorded for analysis.

4.2 Experiment 3

4.2.1 Participants

Forty-eight native speakers of English (sixteen male) with normal reading ability were recruited for this experiment. They were paid £3 for their participation. A typical session lasts about 25 min.

4.2.2 Stimuli and design

A 2 [*Reporting Style*: direct versus indirect speech] × 2 [*Linguistic Focus*: voice-focused versus voice-unfocused pre-speech descriptions] × 2 [*Gender*: female versus male reported speaker] × 3 [*Emotion*: high versus medium versus low emotional arousal of the reported speaker] [8] design was employed (see examples below).

[8] *Emotion* was a between-item manipulation that could not be counterbalanced within items. Hence the examples for the *high*, *medium* and *low* arousal stories are presented separately from the examples in the table below (in the following text).

		Male reported speaker	Female reported speaker
Voice focused	**Direct speech**	*PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. Ella asked for an extension but Jim looked concerned. His voice turned serious: 'Hmm…**we really need those data in by next month for that conference**.'*	*PhD student Jim was summoned to his supervisor Ella's office to give a report on his current progress. Jim asked for an extension but Ella looked concerned. Her voice turned serious: 'Hmm…**we really need those data in by next month for that conference**.'*
	Indirect speech	*PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. Ella asked for an extension but Jim looked concerned. His voice turned serious, saying that **they really needed those data in by next month for that conference**.*	*PhD student Jim was summoned to his supervisor Ella's office to give a report on his current progress. Jim asked for an extension but Ella looked concerned. Her voice turned serious, saying that **they really needed those data in by next month for that conference**.*
Voice unfocused	**Direct speech**	*PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. Ella asked for an extension but Jim looked concerned. His face turned serious: 'Hmm…**we really need those data in by next month for that conference**.'*	*PhD student Jim was summoned to his supervisor Ella's office to give a report on his current progress. Jim asked for an extension but Ella looked concerned. Her face turned serious: 'Hmm…**we really need those data in by next month for that conference**.'*
	Indirect speech	*PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. Ella asked for an extension but Jim looked concerned. His face turned serious, saying that **they really needed those data in by next month for that conference**.'*	*PhD student Jim was summoned to his supervisor Ella's office to give a report on his current progress. Jim asked for an extension but Ella looked concerned. Her face turned serious, saying that **they really needed those data in by next month for that conference**.'*

Sixty octets of short fictitious stories containing different protagonists were prepared as reading materials (The full stimulus set can be found in Appendix 2). Each story started with two declarative sentences to set up a scenario, followed by a speech quotation from a main protagonist in the story. The quotation employed either the direct speech or the indirect speech reporting style (*Reporting Style* factor); each was preceded by either a voice-focused description or a voice-unfocused description (*Linguistic Focus* factor). The reported speaker's gender was either a male or a female speaker (*Gender* factor). The reported speaker's emotional arousal in each item was rated by a different sample of subjects (see 4.2.4.1) and the sixty items were evenly split into 20 'high', 20 'medium', and 20 'low' emotional arousal groups (*Emotion* factor). The following are three examples of the high, medium and low arousal stories:

> [HIGH] *Millionaire Joseph was addicted to betting on horses. Tipped by a so-called 'insider', he recently placed an enormous bet, but shockingly, the horse had lost. Angry with his informant, Joseph shouted furiously on the phone:* ***'Where did you bloody information come from!? That was a huge amount of money – almost one million pounds!'***

[MEDIUM] *Britney is a student at the University of Glasgow. After a heavy snow in the afternoon, she was complaining to her boyfriend James about the weather on their way home. Her voice sounded very grumpy and unpleasant:* ***'I really hate the winter! It's always dark and the roads are too slippery.'***

[LOW] *Smith was working in a small antiques shop down the local high street. Today, a middle-aged posh lady with thick glasses came into the shop. She looked around and said in a nonchalant tone:* ***'You may be surprised to learn that I'm a world-renowned collector of rare memorabilia of White-eared Pheasant.'***

Eight lists of stimuli were created, each containing counterbalanced item-condition combinations. Each item appeared once per list, but in a different condition across lists. The presentation order of the items was randomised for each list.

4.2.3 Procedure

The experiment took place in a sound-attenuated recording room (Voice Neurocognition Lab, CCNi). The participants were given one of the eight stimulus lists. They were instructed to read the stories aloud as naturally and fluently as possible, and only for once; such an instruction did not explicitly encourage participants to vocally 'act out' the stories during reading. After setting up the microphone and the digital recorder, the experimenter sealed the room and the participant started to read the stories aloud while audio recordings were taken using Audacity (http://audacity.sourceforge.net/). The recordings were saved as .WAV files for later analysis.

4.2.4 Data analyses

4.2.4.1 Coding the reported speaker's emotional arousal

A different sample of 9 native speakers of English (3 male) were recruited to rate the reported speaker's emotional arousal for each story. They were presented with the whole list of 60 stories. The stories only contained direct speech quotations without the *Linguistic Focus* manipulation (the voice-focused or –unfocused descriptions were all replaced with *He/She said*; see the example below).

*PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. Ella asked for an extension but Jim looked concerned. (**Context**)*
*He said: 'Hmm…we really need those data in by next month for that conference.' (**Direct-speech**)*

The participants were instructed to read each story carefully and then rate, taking into account the described situation, how excited/agitated they think the reported speaker would be while he/she was speaking (the direct speech quotations were underscored in the actual stimuli), using a 5-point scale: 0 meant 'very calm or sad', 4 meant 'very excited or agitated'. They wrote down the ratings on rating sheets; the latter were collected by the experimenter for analysis.

The collected emotional arousal ratings were averaged by item and were then sorted in an ascending order. The 20 items with the lowest mean ratings were allocated into the 'low' emotional arousal group (M = 1.01, SD = 0.32) while the middle 20 items into the 'medium' emotional arousal group (M = 1.95, SD = 0.34) and the final 20 items with the highest average ratings into the 'high' emotional arousal group (M = 3.05, SD = 0.43).

4.2.4.2 Acoustic analyses of $\overline{F}_0$ on the recordings of the critical quotations

Acoustic analyses of pitch (F_0) were conducted in Praat software (Boersma & Weenink, 2010). Advised by the Praat manual (http://www.fon.hum.uva.nl/praat/manual/), a frequency range of 75 to 300 Hz was applied for male participants and a range of 100 to 500 Hz for female participants.

The audio samples of the critical direct and indirect speech quotations were listened to and manually selected by the experimenter. The F_0 values on the selected samples were extracted by frame (i.e., by roughly every 7.5 ms), and were then averaged ($\overline{F}_0$) for each experimental trial. The $\overline{F}_0$ measures were not suitable for ANOVAs because they were distributed bimodally due to the systematic pitch difference between the male participants (M = 118 Hz) and the female participants (M = 205 Hz). To deal with this problem, they were 'centred' by subtracting their by-subject means such that the 'centred' pitch means (labelled as $\overline{F}_0$') were normally distributed with a grand mean of 0 Hz. Next, trials with a $\overline{F}_0$' of more than 3 SD above a condition's mean were considered as outliers and were removed. In total, 37 trials (1.3% of the trials, equally distributed across conditions) were excluded. The $\overline{F}_0$'s in the remaining 2843 trials were submitted to 2 [*Reporting style*: direct versus indirect speech] × 3 [*Emotion (arousal)*: low versus medium versus high emotional arousal of reported speakers] × 2 [*Gender*: male versus female reported speakers]

× 2 [*Linguistic Focus*: voice-focused versus voice-unfocused pre-speech descriptions] ANOVAs with subject (F_1) and item (F_2) as random factors.

4.2.5 Results and discussion

There was a significant *Reporting Style* main effect, $F_1(1, 47) = 117.588$, $p < .001$, $F_2(1, 59) = 217.113$, $p < .001$. It was found that participants significantly increased their pitch (F_0) during oral reading of direct speech (M = 4.71 Hz, SD = 12.63 Hz) as opposed to indirect speech (M = -5.21 Hz, SD = 8.68 Hz). Although this finding was not central to our hypothesis, it was nonetheless in line with the 'inner voice' experience we had observed in our previous studies; the latter showed that individuals are more likely to mentally simulate vivid vocal depictions of reported speakers during comprehension of direct speech as opposed to indirect speech (Chapter 2 & 3). Likewise, mental simulations of vocal depictions should also take place during oral reading (since it requires comprehension of the written text) of direct speech rather than indirect speech. Thus, the simulated representations of vivid vocal depictions would then be orally interpreted in the form of vivid *voice acting* of the reported speech acts via articulation. Such voice acting in oral reading of direct speech was confirmed both by the experimenter's personal observation of the recorded audio samples and by the significantly larger variation of pitch in the direct speech condition than in the indirect speech condition (quantified in pitch standard deviations over time, SD = 12.63 versus 8.68, paired-sample $t(47) = 11.35$, $p < .001$, $t(59) = 12.61$, $p < .001$). Voice acting requires relatively stronger and more complex vocal modulations of acoustics. To achieve this, oral readers have to increase their lung volume so as well as the subglottal pressure to prepare extra vocal energy. The changes in the subglottal pressure enhance the frequency of vocal fold vibration, thereby increasing the overall F_0. In this sense, the significant *Reporting Style* main effect could have been engendered by engaging in voice acting during oral reading of direct speech rather than indirect speech. Another plausible contributor to this main effect is the F_0 resetting[9] in oral reading of direct speech but not indirect speech. In speech communication, F_0 tends to decline throughout the course of a declarative sentence (Bolinger, 1964; Cohen, Collier, & Thart, 1982; Cohen & t' Hart, 1967). A single F_0 declination could either span an entire

[9] F_0 resetting refers to the phenomenon where the declining F_0 contour jumps up to a relatively higher F_0 level such that the declination preceding and following the resetting can be better predicted by two separate functions rather than one (see an example in Figure 9)

sentence (e.g., Figure 10) or be segmented into multiple 'breath groups' (e.g., Figure 9). Speakers replenish air between these 'breath groups', each starting with a F_0 resetting. In communication, F_0 declination and resetting encode important pragmatic information (Danly et al., 1983; Hird & Kirsner, 2002; Menn & Boyce, 1982; Umeda, 1981; Vaissiere, 1983) such as the 'units of meaning' which speakers plan to convey (which can span from a single word to seven or more clauses (Winkworth, Davis, Adams, & Ellis, 1995). It was observed that direct and indirect speech, as different 'units of meaning' or 'pragmatic entities', were indeed characterised in differential breathing patterns during oral reading. After reading the main clauses of direct and indirect speech (e.g., *He said*), oral readers frequently paused before reading direct speech quotations but rarely do so before reading indirect speech quotations. The pauses before reading direct speech quotations perhaps signal a perspective switch from the reporter to the reported speaker. Nevertheless, such a pause introduces a F_0 resetting and increases the $\overline{F}_0$ for oral reading of the following direct speech quotations (Figure 11). By contrast, without such a pause before oral reading of indirect speech quotations, the F_0 continues to decline from reading of the main clauses, resulting in a relatively lower $\overline{F}_0$ for oral reading of the following indirect speech quotations (Figure 11). Thus, the significant *Reporting Style* main effect may have also been engendered by the differential breathing patterns during oral reading of direct speech versus indirect speech. That said, since we were interested in whether the *Reporting Style* factor interacted with other Context factors (i.e., *Emotion* and/or *Gender*), we did not further validate or distinguish the above two interpretations of the *Reporting Style* main effect.

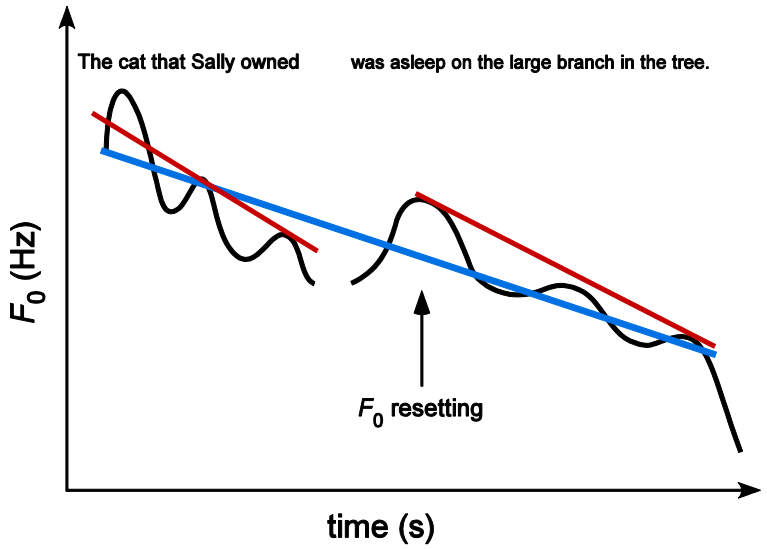


Figure 9 Fundamental frequency (F_0) contour (black) for sentence *The cat that Sally owned was asleep on the large branch in the tree*. The **blue** line represents the prediction for a single F_0 declination function whereas the **red** lines represent the predictions for two declination functions. (Modified from Danly et al., 1983)

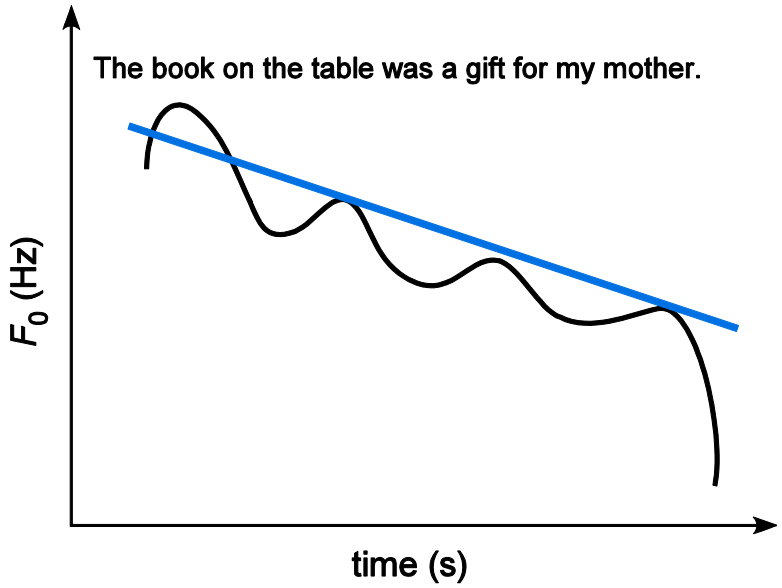


Figure 10 Fundamental frequency (F_0) contour (black) for sentence The book on the table was a gift for my mother. The **blue** line represents the prediction for a single F_0 declination function. (Modified from Danly et al., 1983)

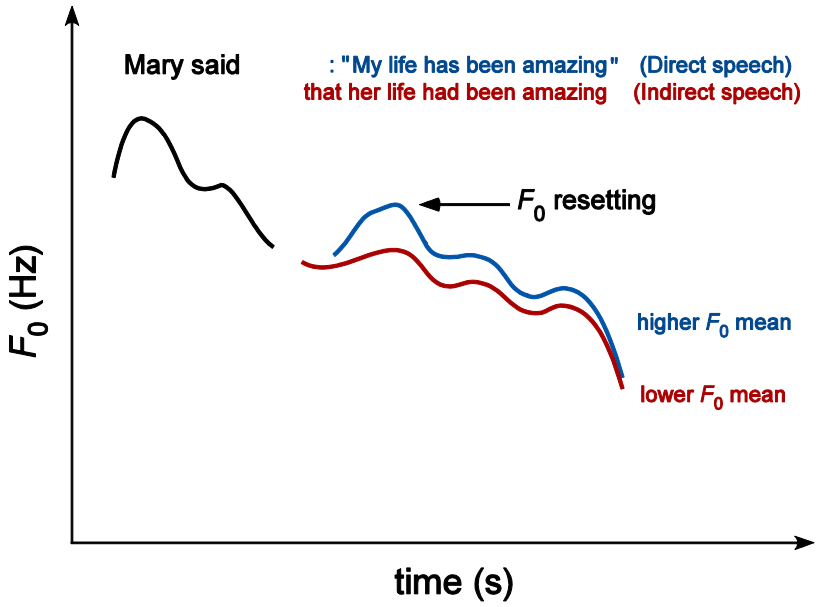


Figure 11 Hypothetical pitch (F_0) contours (black) for direct speech *Mary said: 'My life has been amazing' (blue)* and indirect speech *Mary said that her life had been amazing (red)*. Due to the insertion of a F_0 resetting, oral reading of direct speech elicits higher F_0 on average than indirect speech.

In line with our predictions, we found that the *Reporting Style* factor significantly interacted with the *Emotion* factor, $F_1(2, 46) = 40.478$,.$p < .001$, $F_2(2, 58) = 11.789$, $p < .001$, and the *Gender* factor, $F_1(1, 47) = 6.236$, $p = .016$, $F_2(1, 59) = 9.501$, $p = .003$, respectively. Exploring the former, we found that in the direct speech condition, there was a consistent trend of increasing $\overline{F}_0$ as the reported speaker's emotional arousal increased. The positive relationship between $\overline{F}_0$ and emotional arousal was consistent with the acoustic profiles of vocal emotion expressions (Banse & Scherer, 1996). However, such an increasing trend of $\overline{F}_0$ in relation to the reported speaker's arousal level was not observed during oral reading of indirect speech (Figure 12). Exploring the *Reporting Style* × *Gender* interaction, we found that oral readers significantly increased their pitch when the reported speaker was female as opposed to when the reported speaker was male, but only during oral reading of direct speech and not indirect speech (Table 3). The findings indicate that individuals are more likely to adjust their own pitch in accordance with the reported speaker's emotional arousal and gender during oral reading of direct as opposed to indirect speech. It supports the observation that individuals try to simulate the pitch of a reported speaker's voice during oral reading of direct speech as opposed to indirect speech. That said, further evidence is needed to verify that the underlying representations of the observed interactions are voice-related in nature.

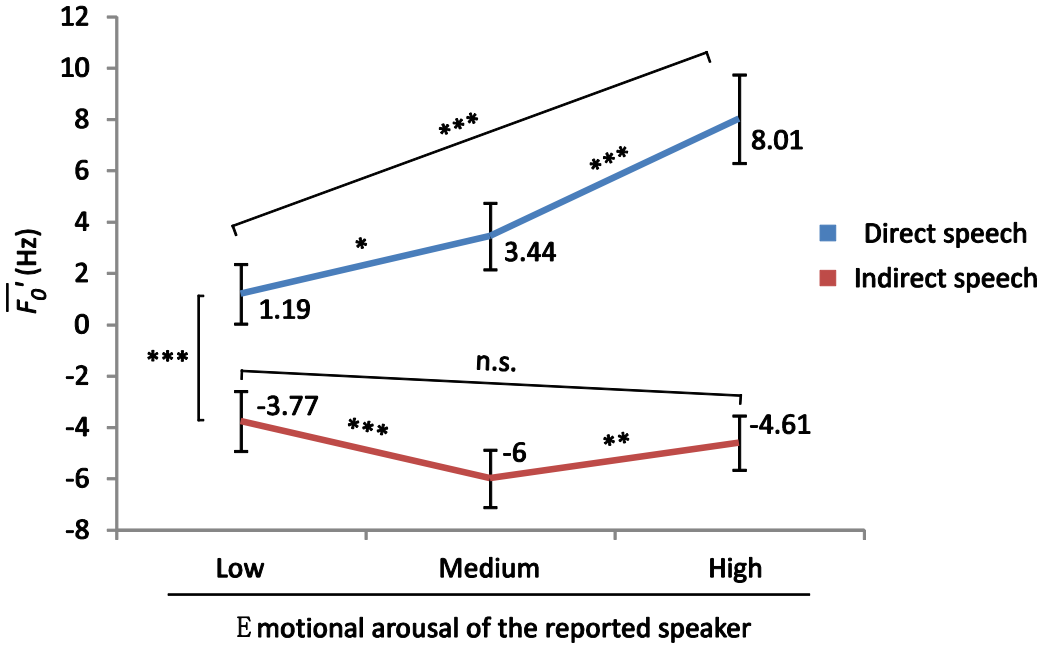


Figure 12 Illustrations of the *Reporting Style* × *Emotion* interaction. The numbers indicate the condition means (in the centred pitch $\overline{F_0}'$). The error bars represent the 95% confidence intervals for the means. The simple effects (both by subjects and by items) of *Emotion* are also shown (n.s. = not significant; * = $p < .05$; ** = $p < .01$; *** = $p < .001$).

			Direct speech	Indirect speech	Reporting Style Contrast			
					By subject		By item	
					t(47)	*p*	*t*(59)	*p*
Female reported speaker			5.28 (14.00)	-5.04 (9.10)	**10.971**	**.000**	**11.764**	**.000**
Male reported speaker			3.14 (13.15)	-4.55 (9.05)	**8.653**	**.000**	**10.823**	**.000**
Gender Contrast	**By subject**	***t*(47)**	1.796	-1.054	Note: Significant contrasts are highlighted in bold.			
		p	.079	.297				
	By item	***t*(59)**	**3.209**	-.833				
		p	**.002**	.408				

Table 3 The means of the centred pitch ($\overline{F_0}'$, in Hz, with standard deviations in parentheses) for the direct versus indirect speech critical quotations across the *Gender* contexts. Also shown are results from 2-tailed sample *t*-tests examining the simple effects of *Reporting Style* and *Gender*.

To this end, we examined the effects of *Linguistic Focus* on these two interactions. We found that the *voice-focused* manipulation had no noticeable effect on the *Reporting Style* × *Emotion* interaction; the three-way interaction was not significant, *F*s < 1.6, *p*s > .2. The results could mean that the underlying mental representations of the *Reporting Style* × *Emotion* interaction may not be voice-related in nature, e.g., it could have been engendered by supra-modal representations of emotional arousal; however, it was also possible that the *Reporting Style* × *Emotion* interaction was too robust to be affected by the *Linguistic Focus* manipulation. Future studies would be needed to disentangle between these two

possibilities. By comparison, the *voice-focused* manipulation appeared to effectively mask the *Reporting Style* × *Gender* interaction: the three-way interaction was significant by item, $F_2(1, 59) = 4.103$, $p = .047$, and was approaching significance by subject, $F_1(1, 47) = 2.719$, $p = .106$[10]. We explored this three-way interaction by splitting our data by *Linguistic Focus* – one data set contained quotations that were preceded with *voice-focused* descriptions (e.g., *Her voice turned serious*) while the other contained quotations that were preceded with *voice-unfocused* descriptions (e.g., *Her face turned serious*). Next, we performed the same 2 *Reporting Style* × 2 *Gender* ANOVAs on these two data sets, respectively. It was discovered that in the *voice-unfocused* data set, the *Reporting Style* × *Gender* interaction was significant, $F_1(1, 47) = 6.624$, $p = .013$, $F_2(1, 59) = 12.914$, $p = .001$. Consistent with our main findings, it was found that oral readers significantly increased their pitch when the reported speaker was female as opposed to when the reported speaker was male, but only during oral reading of direct speech rather than indirect speech[11] (see the upper panel in Table 4). By contrast, in the *voice-focused* data set, the *Reporting Style* × *Gender* interaction was not significant, Fs < .83, ps > .36; in both the direct speech and indirect speech conditions, oral readers' pitch did not differ as a function of the reported speaker's gender (see the lower panel in Table 4). The results appear to be in line with the masking hypothesis (see 4.1 Introduction and Figure 8B). They indicate that linguistic focus on voices can effectively 'mask' the *Gender*-concordant $\overline{F}_0$ modulations during oral reading of direct speech as opposed to indirect speech whereas the linguistic focus on non-voice aspects cannot. Such a 'voice-selective' masking effect suggests that the underlying mental representations of the *Reporting Style* × *Gender* interaction are likely to be voice-related in nature, i.e., contain representations of pitch information.

[10] The three-way interaction was not significant by subject. This was perhaps due to the relatively large variability within a relatively small sample size (48 subjects in relation to a 2 × 2 × 2 design result in mere 6 data points per item per condition), suggesting a lack of statistical power rather than a null effect.

[11] The by-subject *Gender* contrast was marginally significant due to the relatively large variability within a relatively small sample size (also see Footnote 10).

Voice Unfocused data set			Direct speech	Indirect speech	Reporting Style Contrast			
					By subject		By item	
					t(47)	*p*	*t*(59)	*p*
Female speaker (*Her face turned serious*)			5.73 (14.16)	-5.87 (8.83)	**9.005**	**.000**	**9.872**	**.000**
Male speaker (*His face turned serious*)			2.85 (12.69)	-4.65 (8.34)	**6.298**	**.000**	**8.545**	**.000**
Gender Contrast	By subject	*t*(47)	1.963	-1.744	Note: Significant contrasts are highlighted in bold text			
		p	.056	.088				
	By item	*t*(59)	**3.050**	-1.667				
		p	**.003**	.101				
Voice Focused date set			Direct speech	Indirect speech	Reporting Style Contrast			
					By subject		By item	
					t(47)	*p*	*t*(59)	*p*
Female speaker (*Her voice turned serious*)			5.88 (13.87)	-5.01 (9.33)	**9.670**	**.000**	**10.286**	**.000**
Male speaker (*His voice turned serious*)			4.74 (13.55)	-5.29 (9.72)	**8.618**	**.000**	**10.002**	**.000**
Gender Contrast	By subject	*t*(47)	.909	.214	Note: Significant contrasts are highlighted in bold text			
		p	.368	.832				
	By item	*t*(59)	1.300	.386				
		p	.199	.701				

Table 4 The means of the centred pitch ($\overline{F}_0$', in Hz, with standard deviations in parentheses) for the direct versus indirect speech critical quotations across the *Gender* contexts in the two data sets separated by *Linguistic Focus*. Also shown are results from 2-tailed sample *t*-tests examining the simple effects of *Reporting Style* and *Gender* in each data set.

In summary, we have demonstrated that oral readers spontaneously adjust their pitch in accordance with the linguistically implied reported speaker's emotional arousal and gender during reading aloud of direct speech but not (or less so) during reading aloud of indirect speech. The *Gender*-concordant pitch modulations during oral reading of direct as opposed to indirect speech can be selectively masked by exerting linguistic focus on the reported speaker's voices, indicating that the underlying representations are likely to contain acoustic information of pitch. The results are in line with our previous findings (Chapter 2 & 3), suggesting that individuals are more likely to mentally simulate voice representations during comprehension of direct as opposed to indirect speech. More specifically, we show that pitch may be an acoustic aspect of such voice representations, at least in relation to reported speakers' genders, and that mental simulations of voices are not only reflected in brain-activation patterns, but also in modulations of articulation during oral reading.

4.3 Experiments 4, 5 and 6

The following three experiments were carried out to investigate mental simulations of pitch during *silent reading* of direct versus indirect speech, i.e., without the demand for articulation. We did not investigate the effects of reported speakers' emotional arousal because linguistic focus on voices did not have a noticeable effect on *Emotion*-concordant pitch modulations during oral reading of direct versus indirect speech in Experiment 3. This suggests that the *Emotion* effects on pitch modulations may be representationally voice-unrelated. For this reason, we *exclusively* focused on the effects of reported speakers' genders on mental simulations of pitch representations in the following investigations; this was based on the findings in Experiment 3 that the representations underlying the observed *Gender*-concordant pitch modulations during oral reading of direct rather than indirect speech are likely to be voice-related in nature since they could be selectively masked by exerting linguistic focus on voices.

4.3.1 Participants

Twenty-four native speakers of English participated in each of the three experiments in exchange for £5. They had no reported reading or hearing disabilities. A typical experimental session took approximately 45 min.

4.3.2 Stimuli and design

A 2 [*Reporting Style*: direct versus indirect speech] × 2 [*Gender Match*: the reported speaker's gender matches versus mismatches the probe word speaker's gender] design was used in all three experiments. Sixty pairs of written stories with different protagonists (indicated by different English names) were prepared as reading materials (see examples below). They were almost the same as in Experiment 3, apart from that the *Linguistic Focus* manipulation was removed and the *Gender* manipulation was now implemented between items for the purpose of design simplification.

	Gender Match	Gender Mismatch
Direct speech	*PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. Ella asked for an extension but Jim looked concerned.* ***He*** *said: 'Hmm, we really need those data in by next month for that conference.'* PROBE WORD: *'conference',* ***male*** *voice*	*PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. Ella asked for an extension but Jim looked concerned.* ***He*** *said: 'Hmm, we really need those data in by next month for that conference.'* PROBE WORD: *'conference',* ***female*** *voice*
Indirect speech	*PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. Ella asked for an extension but Jim looked concerned.* ***He*** *said that they really needed those data in by next month for that conference.* PROBE WORD: *'conference',* ***male*** *voice*	*PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. Ella asked for an extension but Jim looked concerned.* ***He*** *said that they really needed those data in by next month for that conference.* PROBE WORD: *'conference',* ***female*** *voice*

Each of the pair stories consisted of two declarative sentences to set up a scenario, followed by either a direct speech or an indirect speech quotation (*Reporting Style* factor). Half of the 60 items described utterances reported from female speakers while the other half from male speakers. Spoken words were recorded as auditory probes for each item. They were recorded from either a female speaker or a male speaker such that they either matched or mismatched the reported speaker's gender indicated in the story (*Gender Match* factor). To ensure that the male-speaking and the female-speaking spoken probe words differed only in pitch rather than other acoustics, we normalised the intensity (dB) of all probe word recordings using MatLab scripts (provided by the Voice Neurocognition Laboratory, University of Glasgow), and adjusted the durations of the recording files (i.e., the lengths of the spoken words) and kept them proximate within each item. Acoustic analyses on the edited probe words (see Table 5) confirmed that the male-speaking words and the female-speaking words differed from each other only in pitch. While the male-speaking probe words were significantly lower in pitch than the female-speaking probe words (M_{diff} = 74.6 Hz), there were no meaningful/significant differences between them in loudness or duration; although the male-speaking words were significantly louder than the female-speaking words, the difference amounted to a mere 0.07 dB (in loudness perception given that a 5 dB increment step size is usually recommended for subjects with normal-hearing sensitivity, Cox, Alexander, Taylor, & Gray, 1997). This indicates that our intended manipulation of the pitch of the spoken words was effective. In addition to the acoustic manipulations, we also controlled the presentation position of the corresponding critical words in text to keep the presentations of the critical words and the spoken probe words temporally close: We always selected, within a given speech quotation, the final word as the critical word (e.g., the spoken word *conference* is the last word in *He said:*

'Hmm, we really need those data in by next month for that conference.'); this was to ensure that the spoken probe word would be presented soon enough after reading the critical word. Also prepared were yes/no comprehension questions regarding the contents of the stories (e.g., *Was Ella Jim's PhD student?*) which would assess participants' overall comprehension performance and to ensure that they read the stories for comprehension.

	Male-spoken probe words	Female-spoken probe words	Paired-sample t-tests by item	
			***t*(59)**	***p***
Pitch (F_0) (Hz)	110.8 (13.4)	185.4 (17.9)	-29.877	<.001
Loudness (dB)	78.46 (0.35)	78.39 (0.25)	2.080	.042
Duration (ms)	674 (119)	678 (118)	1.628	.109

Table 5 Means for pitch, loudness and durations (along with standard deviations in parentheses) on the audio recordings of the male-speaking and the female-speaking probe words. Also shown are results for the paired-sample *t*-tests investigating the simple contrasts of the *Probe Word Speaker's Gender*.

In addition to the 60 critical items, 120 filler stories were prepared to conceal the intended experimental manipulations. The filler stories were mostly declarative narratives. Twenty-eight of them (ca. 23%) contained direct speech (14 items) or indirect speech sentences (14 items) but without the critical experimental manipulations. In Experiment 4 and 6, one third of the filler items (40 items) contained spoken probe words which *matched* the critical words in the written stories. The remaining 80 filler items contained *mismatch* spoken words which were synonyms to the relevant critical words. Together there were 100 *word match* items (60 critical items + 40 fillers) and 80 *word mismatch* items (fillers). In Experiment 5 (lexical decision task), one third of the filler items (40 items) contained English meaningful words which either *matched* or *mismatched* the critical words in the written stories. The remaining 80 filler items contained non-words that were digitally constructed using the recordings of the words. For example, the non-word "bendarving" was created using the recordings of "banana" and "starving". Overall, there were 100 English meaningful word items and 80 non-word items.

Four lists of counterbalanced item-condition combinations were generated. A given trial could occur in one of the four conditions: direct speech – male probe word, indirect speech – male probe word, direct speech – female probe word, and indirect speech – female probe word. Each item appeared in one condition per list and in different conditions across lists.

Each participant only saw one list. The presentation order was randomised for each participant.

4.3.3 Procedure

The experiment was conducted in a lab equipped with a Dual Core Dell PC. Participants were seated in a comfortable chair, wearing headphones. Stimuli were presented using E-prime 2.0 presentation software (Psychology Software Tools, Inc., USA).

4.3.3.1 Procedure for Experiment 4

After a ten-trial practice, participants were presented with the given stimulus list in the following manner. In each trial, participants first saw a cue 'NEW TRIAL' at the centre of the computer screen for 1000 ms, which was followed by the presentation of a short story. The text was presented in a 15 pt Arial font on a light grey background. The whole text passage was centred and wrapped within 50% width of the screen, together making the reading stimuli appear as natural as possible. Participants were instructed to silently read the text passage in their own time, and press spacebar to terminate the presentation when they finished reading. Next, they saw a question mark at the centre of the screen for 1000 ms and then heard a spoken word through the headphones. Participants were required to indicate whether the spoken word had been mentioned in the story they had just read as quickly and accurately as possible by pressing the left mouse button for 'yes' or the right one for 'no'. Their reaction times (RTs) and recognition accuracy on the spoken words and were recorded. Finally, in one third of the trials, they were required to answer a yes/no comprehension question regarding the story they had just read. The answering accuracy was also recorded.

4.3.3.2 Procedure for Experiment 5

The procedure was the same as Experiment 4 except for that in Experiment 5, a lexical decision task was employed: As participants heard a spoken probe word via the headphones, they had to indicate, as quickly and accurately as possible, whether the heard spoken word was an English meaningful word or not, by pressing the left mouse button for 'yes' or the right mouse button for 'no'.

4.3.3.3 Procedure for Experiment 6

Instead of using self-paced presentation of text for silent reading, Experiment 6 employed serial presentation of text. The reading materials were segmented and presented in serial. For example, direct speech sentence *He said: 'Hmm, we really need those data in by next month for that conference.'* was segmented (the boundaries are indicated by '//') into *He said: '//Hmm,//we really need// those data in// by next month// for// that// conference.'* The presentation duration of each segment was determined by 100 ms + 50 ms × number of syllables + 200 ms × number of words, allowing sufficient time for reading. Critically, the presentation would pause on the critical word (e.g., the final word 'conference') while a spoken probe word would be presented *simultaneously* through the headphones. Participants had to respond as quickly and accurately as possible, using the mouse button (left for 'yes', right for 'no'), as to whether the word that they had read and the word that they had heard were the same or not.

4.3.4 Results and discussion

Incorrect responses to the spoken words were discarded (2.5% of the trials in Experiment 4, 0.7% and 2.3% in Experiment 5 and 6, respectively). Outliers were excluded in two steps: First, the correct responses with a RT of longer than 2000 ms or shorter than 250 ms were removed. Next, the remaining trials with a RT which was 3SD away from a condition's mean were excluded. In total, 5.5% of the trials were excluded from analysis in Experiment 4, 2.5% and 4.7% in Experiment 5 and 6, respectively.

The remaining trials were submitted to 2 [*Reporting Style*: direct versus indirect] × 2 [*Gender Match*: the probe word's gender matched versus mismatched the reported speaker's gender] ANOVAs with subject (F_1) and item (F_2) as random factors, respectively.

The accuracy on comprehension questions was above 70% across all three experiments (Experiment 4: M = 86.7%, SD = 6.2%; Experiment 5: M = 87%, SD = 7.8%; Experiment 6: M = 86.8%, SD = 7.6%), indicating that subjects comprehended the reading materials attentively to a reasonable extent.

We did not find any significant effect or interaction in any of three experiments, except that in Experiment 6 (using the *immediate probing* paradigm where the written critical

word and the spoken probe word were presented *simultaneously*), the *Reporting Style* × *Gender Match* interaction was approaching significance, $F_1(1, 23) = 3.369$, $p = .079$, $F_2(1, 59) = 2.454$, $p = .123$. Exploring this interaction, it was found that in the direct speech condition, participants' response times (RTs) to the spoken probe words were faster when the probe word speaker's gender matched the reported speaker's gender ($M = 742$ ms) as opposed to when they mismatched ($M = 760$ ms). Such a match *advantage* in RTs is in line with mental simulations of pitch representations during silent reading of direct speech: The recognition of the gender-match spoken words was facilitated as compared to the recognition of the gender-mismatch spoken words because the former contained pitch information that was more congruent with the mentally simulated pitch representations during silent reading. By contrast, in the indirect speech condition, the RT difference across the *Gender Match* manipulation in the indirect speech condition was in the opposite direction (744 ms versus 736 ms for the match versus mismatch conditions). This match *disadvantage* was inconsistent with the pitch simulation hypothesis and amounted to only half as much as compared to the direct speech condition (the RT difference = 8 ms in indirect speech versus 18 ms in direct speech). Such result patterns seemed to resemble a gender-match advantage in silent reading of direct speech and a null effect in silent reading of indirect speech, which was in line with our prediction. Thus, we suspected that there may have been latent factors which may have diminished the significance of the *Reporting Style* × *Gender Match* interaction.

Indeed, in post-hoc analyses of Experiment 6 (immediate probing), we found that the *Reporting Style* × *Gender Match* interaction seemed to interact with the *Reported Speaker's Gender* (the three-way interaction was approaching significance, $F_1(1, 23) = 3.347$, $p = .080$, $F_2(1, 59) = 1.863$, $p = .178$). We separated the data into two groups: One group consisted of items with female reported speakers while the other with male reported speakers. We then performed the same *Reporting Style* × *Gender Match* ANOVAs on the two sets of data, respectively. We found that when the reported speaker was female, the *Reporting Style* × *Gender Match* interaction was significant both by subject and by item, $F_1(1, 23) = 4.890$, $p = .037$, $F_2(1, 59) = 5.036$, $p = .032$, whereas when the reported speaker was male, this interaction was not significant, Fs < .14, ps > .7. Exploring the former, we found that in silent reading of direct speech, participants' RTs to the spoken probe words were significantly faster when the probe word speaker's gender matched the reported

speaker's gender than when they mismatched; whereas in silent reading of indirect speech, such *Gender Match* effects were not observed (Table 6).

Female-speaking items			**Direct speech**	**Indirect speech**	**Reporting Style Contrast**			
					By subject		**By item**	
					***t*(23)**	***p***	***t*(29)**	***p***
Gender Match			741 (192)	759 (194)	-1.381	.181	-.800	.430
Gender Mismatch			782 (242)	745 (217)	1.718	.099	2.036	.051
Gender Match Contrast	**By subject**	***t*(23)**	**-2.524**	.826	Note: significant contrasts are highlighted in bold text			
		p	**.019**	.418				
	By item	***t*(29)**	-2.030	.630				
		p	.05	.534				
Male-speaking items			**Direct speech**	**Indirect speech**	**Reporting Style Contrast**			
					By subject		**By item**	
					***t*(23)**	***p***	***t*(29)**	***p***
Gender Match			741 (180)	732 (188)	.402	.692	.317	.753
Gender Mismatch			738 (204)	724 (203)	.695	.494	.595	.556
Gender Match Contrast	**By subject**	***t*(23)**	.237	.937	Note: significant contrasts are highlighted in bold text			
		p	.815	.358				
	By item	***t*(29)**	-.022	.172				
		p	.983	.865				

Table 6 Reaction times (in ms, with standard deviations in parentheses) for the direct versus indirect speech conditions across the *Gender Match* manipulations in the two data sets separated by the gender of the reported speaker (Experiment 6). Also shown are results from 2-tailed sample *t*-tests examining the simple effects of *Reporting Style* and *Gender Match* in each data set.

The results show that individuals are more likely to mentally simulate voice representations (which contain pitch information) in accordance with the reported speaker's gender during silent reading of direct as opposed to indirect speech, but only when these utterances are reported from a female speaker. Such 'female-favoured' voice simulations could perhaps be explained by women's superior expressivity over men during verbal communication (e.g., Kring & Gordon, 1998; Robinson & Johnson, 1997): Female speakers are more likely to employ vivid vocal depictions in verbal communication than their male counterparts. Thus, comprehension of female-speaking direct speech should be grounded in mental simulations of more vivid vocal depictions that are employed by female speakers whereas comprehension of male-speaking direct speech would be grounded in experiences of not-so-vivid utterances by male speakers. This could explain why individuals are more likely to simulate voices when reading female-speaking direct speech as opposed to male-speaking direct speech. Although the three-way (*Reported Speaker's* × *Gender Reporting Style* × *Gender Match*) interaction did not reach .05

significance, the post-hoc analyses nonetheless provide a potential hint as to how experiences with differential expressivities between male and female speakers underlie language comprehension. To further validate this 'gender-dependent' voice simulation phenomenon, future research will need to replicate the current results with more strictly counterbalanced design and more carefully manipulated stimuli. For example, the reported speaker's gender may be manipulated within each item to rule out the potential item-specific factors that may have been confounded with the *reported speaker's gender* factor (for design simplification, the current investigations merely manipulated the reported speakers' gender between items).

It should be noted that this gender-dependent voice simulation phenomenon in Experiment 6 (immediate probing) was the only effect that we observed in all three experiments. It may reflect that the simulated voice representations during silent reading of direct rather than indirect speech are 'short-lived', similar to mental simulations of actions (Kaschak & Borreggine, 2008). This could perhaps explain why there was no effect detected in Experiment 4 or 5: The 'off-line' paradigms we employed in Experiments 4 and 5 present spoken probe words 1000 ms after the offset of silent reading; this setup allows 'short-lived' voice representations to 'decay' before the presentation of the spoken probe words, and it is therefore not sensitive enough to detect such representations. By contrast, the 'on-line' paradigm we employed in Experiment 6 presents the written critical word and the spoken probe words simultaneously; this setup does not allow the target representations to 'decay', and therefore it was sensitive enough to detect the gender-dependent voice simulations. Although this higher detecting sensitivity of 'on-line' paradigms was no more than speculations, we nevertheless believed it would be sensible to exclusively focus on 'on-line' paradigms for our following investigations (e.g., Chapter 5 & 6).

In sum, Experiments 4 - 6 investigated mental simulations of pitch information during silent reading of direct versus indirect speech. We examined the effects of the reported speaker's gender on silent readers' recognition latencies to the gender-congruent and gender-incongruent spoken words. Using an immediate probing paradigm in Experiment 6, we found that when the reported speakers are female (e.g., *She said*) participants responded faster to spoken words that are spoken by a same-gender speaker (i.e., a female speaker) than to spoken words that are spoken by an opposite-gender speaker (i.e., a male speaker) during silent reading of direct speech but not indirect speech; when the reported speakers

are male (e.g., *He said*), such *Gender Match* effects were not observed. The results suggest that individuals are more likely to mentally simulate pitch information in accordance with the reported speaker's gender during silent reading of direct as opposed to indirect speech, but only when the reported speaker was female.

4.4 Summary and general discussion

The current investigation made the first attempt at decomposing the 'inner voice' experiences that we had observed during comprehension of direct versus indirect speech into acoustic parameters. Specifically, we examined the pitch aspect of such 'inner voice' representations during both oral reading (Experiment 3) and silent reading (Experiments 4-6) of direct versus indirect speech.

Experiment 3 showed that individuals spontaneously adjusted their own pitch in accordance with the reported speaker's emotional arousal and gender (as derived from the context in which the reported utterances were embedded) during oral reading of direct as opposed to indirect speech. The *gender*-concordant pitch modulations can be masked by exerting linguistic focus on reported speaker's voices, indicating that the underlying mental representations are voice-related in nature and contain pitch information. The findings are in line with our prediction, suggesting that individuals may mentally simulate the reported speaker's pitch representations during oral reading of direct speech but not indirect speech.

Experiments 4 – 6 investigated mental simulations of pitch representations in silent reading of direct versus indirect speech. Participants silently read story-embedded direct and indirect speech and responded to male-speaking or female-speaking spoken words which either matched or mismatched the reported speaker's gender described in the stories. It was found that participants responded faster to the gender-match spoken words than to the gender-mismatch spoken words during silent reading of direct speech rather than indirect speech, but only when the reported speaker was female. The results, although still preliminary, suggest that the acquired experience with the differential expressivities between direct and indirect speech acts as well as the ones between male and female speakers underlies language comprehension.

A previous study on auditory imagery (Abramson, 2007) observed similar gender-concordant 'inner voice' experiences during silent reading. In this study, participants were first familiarised with a male voice and a female voice. They were then instructed to silently read sentences, and had their implicit (auditory lexical decision) or explicit memory (auditory recognition) for words assessed afterwards. In the silently read sentences, the gender of speaker was implied in the initial words, e.g., *'He said,...'* or *'She said,...'*. Auditory lexical decision priming was found for gender-congruent items following silent reading, but only up to 5 min after silent reading. In a second study, similar lexical decision priming was found following listening to sentences, although these effects remained reliable after a 2-day delay. These results suggest that readers create episodic traces of text from auditory images of silently read sentences as they do during listening. The gender-congruency effects observed in this study also indicate the possibility that pitch representations (given its critical role in signalling gender information) might be encoded in silent reading.

That said, although similar, the 'inner voice' phenomenon observed in Abramson's study and the 'inner voice' phenomenon observed in our investigations may reflect distinct cognitive processes. The former may be interpreted as a comprehension-independent, encouraged auditory imagery process whereas the latter is a spontaneous auditory simulation process which is an integral part of language comprehension. In Abramson's study, participants were first familiarised with two voices of different genders before the actual reading trials, followed by cues (e.g.., *He said*) to the identity of the to-be-imagined speaker within the actual reading materials. This type of experimental manipulation (i.e., pre-exposure to specific voices and/or imagery cues during reading) arguably encourages participants to imagine speaker-specific voices during silent reading. Hence, this experimental setup was limited in determining whether such an auditory imagery would also occur as a result of 'normal' language comprehension. By contrast, in our investigations, participants were not presented with any specific voice samples before reading. During reading, they read stories which contained only fictitious protagonists that they did not know. Under this experimental setup, participants were unlikely to imagine any specific, familiar voice. Nevertheless, they still activated voice-like representations during reading, which may be better interpreted in mental simulations or a more spontaneous imagery of unspecific, partial voice representations, as an integral part of

language comprehension. Moreover, our investigations also demonstrate that such an 'inner voice' experience can not only be modulated by linguistic cues to the reported speaker's gender, but can also be modulated by pragmatically distinct reporting styles (i.e., direct versus indirect speech) and linguistic focus (on voices versus on non-voice aspects of the reported speech acts). Thus, our findings may be better interpreted, in line with the demonstration theory of direct speech (Clark & Gerrig, 1990) and the embodied theories of language comprehension (Barsalou, 1999), as mental simulations of gender-concordant pitch information that are guided by language comprehension, rather than as auditory imagery of voices that are encouraged by non-linguistic factors.

In conclusion, our investigations address the pragmatics of different quotation styles and their associated mental representations during text comprehension. Our experiments showed that oral readers spontaneously adjust their pitch to contextually implied pitch when reading direct speech as opposed to meaning-equivalent indirect speech quotations, and that silent readers respond differentially to spoken words that are spoken by same-gender (as the reported speaker) or different-gender speakers but only when reading female-speaking direct speech quotations. These behavioural phenomena (although requiring further validation) may be interpreted as perceptual simulations of reported speakers' voices during comprehension of direct speech rather than indirect speech. Specifically, it suggests that frequency representations such as pitch might be an acoustic aspect of such simulated 'inner voice' experiences during comprehension of direct speech as opposed to indirect speech.

Chapter 5 Contextual modulations of loudness representations in oral and silent reading of direct versus indirect speech

5.1 Introduction

In parallel with our investigations on the pitch aspect in voice simulations, the investigations of the current chapter examined another acoustic aspect in such voice simulation processes, namely the contextually implied *loudness* of the quoted speech. Similar to pitch, loudness is an important acoustic characteristic of a speaker's voice for it signals various types of information about a speaker (e.g., the speaker's emotional state; Banse & Scherer, 1996) and his/her relations to the surrounding environments (e.g., the surrounding noise level and the physical distance between the speaker and the listener). We reasoned that if the simulated 'inner voice' entails acoustic information such as pitch, similar to the voices in the physical world, it is likely that such an 'inner voice' may also encode loudness representations. Thus, we employed the paradigms that we had developed in Chapter 4 and investigated contextually-mediated mental simulations of loudness in both oral reading (Experiment 7) and silent reading (Experiment 8) of direct versus indirect speech.

Experiment 7 investigated whether individuals would adjust their own loudness (measured in mean intensity (dB)) in accordance with the contextually implied loudness of the reported speaker during oral reading of direct as opposed to indirect speech. Participants orally read a series of written stories containing either direct or indirect speech while audio recordings were taken for acoustic analyses. It was predicted that they would spontaneously increase their reading loudness when the preceding linguistic context implied a loud-speaking reported speaker as opposed to a quiet-speaking reported speaker during oral reading of direct speech but not indirect speech. Unlike Experiment 3 which manipulated the reported speaker's pitch level in relation to his/her emotional arousal and gender specifically, Experiment 7 manipulated the reported speaker's loudness level by combining various types of linguistic information including reference to the environmental noise level (e.g., party music), loudness-related verbs (e.g., shouted) and adverbs (e.g., loudly), and the reported speaker's emotional state (e.g., angry). We were not interested in the *individual* effects of each type of linguistic information on reported speakers' loudness; instead, we were only interested in the *joint* effects of these linguistic manipulations of the

reported speaker's loudness level. Thus, for the purpose of design simplification, we collapsed across these linguistic manipulations into a single 'loudness' manipulation; we also dropped Experiment 3's *Linguistic Focus* manipulation[12].

Experiment 8 employed the silent reading paradigm from Experiment 6 (i.e., the immediate probing paradigm in which the written critical word and the spoken probe word are presented simultaneously). We examined the effects of the implied loudness of reported speaker on the recognition latencies to a spoken probe word when the word's loudness was made relevant by the previous text passage. Participants silently read short stories containing direct or indirect speech utterances. The story was segmented first and then presented in a serial fashion. The presentation of direct or indirect speech quotations paused on a critical word while a spoken word was presented simultaneously. Participants had to answer 'yes' or 'no' to indicate whether the spoken word they heard was the same word as the one they saw on the screen. The spoken probe words were acoustically manipulated into either a loud or a quiet version such that it either matched or mismatched the loudness of the reported speaker that was linguistically implied in the story. We acknowledge that the loudness of a speaker's voice may not be solely determined by the absolute volume of a voice. For instance, one can shout quietly or shout from a distance. In these situations, although the voices are low in volume, they would still be perceived as loud. However, it would be impractical to consider possible factors which determine vocal loudness all together in one go. For the purpose of simplification, we only focused on the effects of vocal volumes (still labelled as *loudness*) and manipulated our spoken probe words along this dimension in our current investigation. We predicted that participants' recognition on loudness-congruent spoken words (i.e., the reported speaker's loudness is congruent with the spoken word's loudness) should be facilitated as opposed to recognition on loudness-incongruent spoken words or vice versa.

[12] The exclusion of the *Linguistic Focus* manipulation reduced the required subject number by 50%, which significantly helped facilitate the progress of this research, especially under the tight time schedule. Although we were no longer able to address the acoustic nature of the mental representations that underlie mental loudness simulations in the current investigation, we could still address this question in the silent reading study and in the future fMRI studies. Thus, for the sake of getting preliminary results as quickly as possible, this exclusion was the optimal choice.

5.2 Experiment 7

5.2.1 Participants

Twelve native speakers of English (6 male) with no reported reading impairments participated in exchange for £2. A typical session lasted about 15 minutes.

5.2.2 Stimuli and design

A 2 [*Reporting Style*: direct versus indirect speech] × 2 [*Context*: loud-speaking versus quiet-speaking reported speaker] design was used in this experiment. Twenty-four pairs of short fictitious stories with different protagonists were prepared as reading materials (see below for an example and see Appendix 3 for the full stimulus set). Each story contained either a direct speech or an indirect speech quotation (*Reporting Style* factor). As for each pair, one story described a situation where the protagonist would speak loudly whereas the other a situation where the speaker would speak more quietly (*Context* factor). Crucially, the critical sentence (bold texts in the following example) within each reporting style was identical across contexts and virtually the same across reporting styles.

	Loud-speaking context	**Quiet-speaking context**
Direct speech	*Katy's friends were having such a great time blasting cheesy Christmas music in her flat that she barely heard the knock at the door at 1am. She was expecting another friend and was surprised to find a policeman looking crossly over her shoulder at the party going on inside. He looked very serious, demanding: 'Excuse me Madam. I suggest* ***you keep the noise down as there's been a complaint from your neighbours.****'*	*Katy was reading in the library and had put ear plugs in to block out the other students studying nearby. She absentmindedly began to click her tongue, not realising that other people could hear her until a concerned librarian tapped her lightly on the shoulder. He politely said: 'Excuse me Madam. I suggest* ***you keep the noise down as there's been a complaint from your neighbours.'***
Indirect speech	*Katy's friends were having such a great time blasting cheesy Christmas music in her flat that she barely heard the knock at the door at 1am. She was expecting another friend and was surprised to find a policeman looking crossly over her shoulder at the party going on inside. He looked very serious, demanding that* ***she should keep the noise down as there had been a complaint from her neighbours.***	*Katy was reading in the library and had put ear plugs in to block out the other students studying nearby. She absentmindedly began to click her tongue, not realising that other people could hear her until a concerned librarian tapped her lightly on the shoulder. He politely said to her that* ***she should keep the noise down as there had been a complaint from her neighbours.***

Four stimulus lists with counterbalanced item-condition combinations (6 stories per condition per list) were generated. Each item appeared once per list, but in a different condition across lists. The lists were word-processed into four written scripts. The order

of the stories per script was randomised. Each script was randomly assigned to a quarter of our participants.

5.2.3 Procedure

The experiment was conducted in a sound-attenuated room. Participants were instructed to read out the stories from the script in one go and as naturally and fluently as possible. The instructions did not explicitly encourage participants to vocally 'act out' the stories during reading.

After setting up the microphone and the digital voice recorder, the experimenter sealed the room and the participant started reading the stories aloud while audio recordings were taken using Audacity (http://audacity.sourceforge.net/). The recordings were saved as .WAV files for later analyses.

5.2.4 Results and discussion

Acoustic analyses of intensity (dB) were conducted in Praat (Boersma & Weenink, 2010). The audio recordings of the critical sentences were listened to and manually selected by the experimenter. The intensity values on these recordings were extracted by frame, i.e., by roughly every 7.5 ms. In line with the standards for intensity analysis (according to Praat manual; http://www.fon.hum.uva.nl/praat/manual/), only the intensity values between 50 dB and 100 dB were considered as valid for the analysis. They were first averaged by trials and the by-trial intensity means were then submitted to 2 [*Reporting Style*: direct versus indirect speech] × 2 [*Context*: loud-speaking versus quiet-speaking reported speaker] ANOVAs with subject (F_1) and item (F_2) as random factors.

The *Reporting Style* × *Context* interaction was significant by subject, $F_1(1, 11) = 11.85$, $p = 0.006$, and marginally significant by item, $F_2(1, 23) = 3.805$, $p = 0.063$[13]. Exploring the interaction, we found that direct speech quotations were read out significantly more loudly

[13] The marginally significant interaction by item, as compared with the significant by-subject interaction, suggests that there might have been 'nuisance' items in which the loudness manipulations were not as effective as intended (presumably due to our more unspecific ways of manipulating loudness). Nevertheless, given that the interaction pattern was consistent with our hypothesis and the critical pairwise contrasts were significant as expected, this marginally significant interaction by item may simply reflect the lack of statistical power rather than a null effect.

when the preceding linguistic context implied a loud-speaking as opposed to a quiet-speaking reported speaker. By contrast, such contextual modulations of loudness were not found in oral reading of indirect speech (Table 7). The results showed that our participants adjusted their loudness in accordance with the linguistically implied loudness of the reported speaker during oral reading of direct speech as opposed to indirect speech. The results cannot be explained by differences in phonological or morphological features (tense, person, etc.) between the two *Reporting Style* conditions because the interaction involved differential effects of *Context* within each reporting style (the critical sentences were identical). The current findings are in line with our prediction, suggesting that oral readers may mentally simulate the reported speaker's voice representations during oral reading of direct as opposed to indirect speech. More specifically, we show that loudness may be another acoustic aspect of such voice representations, and that mental simulations of voice are not only reflected in brain-activation patterns, but also in modulations of vocal acoustics during oral reading. However, without the *Linguistic Focus* manipulations (cf. 4.1 Introduction for the rationale), we did not have the relevant information to verify the voice-related nature of the underlying mental representations. Therefore, we would have to address this question in silent reading (Experiment 8) and in our future investigations.

			Direct Speech	Indirect Speech	Reporting Style Contrast			
					By subjects		By items	
					t(11)	*p*	*t*(23)	*p*
Loudly-speaking Context			66.3 (4.5)	63.7 (3.4)	**4.423**	**.001**	**5.147**	**< .001**
Quietly-speaking Context			63.7 (3.9)	63.2 (3.8)	**2.722**	**.02**	.858	.4
Context Contrast	**By subjects**	***t*(11)**	**3.632**	1.497	Note: Significant contrasts are highlighted in bold text			
		p	**.004**	.163				
	By items	***t*(23)**	**3.306**	.631				
		p	**.003**	.534				

Table 7 The intensity means (in dB, along with standard deviations in parentheses) across *Context* on the critical direct versus indirect speech quotations in Experiment 7. Also shown are results from 2-tailed paired sample *t*-tests examining the simple effects of *Context* and *Reporting Style*.

5.3 Experiment 8

5.3.1 Participants

Sixty-four native speakers of English with no reported reading or hearing impairments participated in exchange for £3. A typical session took around 20 minutes. Three participants had to be excluded from analysis because their accuracy on the comprehension questions was below 70%, the latter suggesting that they did not read the story attentively. The remaining sixty-one subjects (41 female, 20 male, age 17 – 37) were valid for the following analyses.

5.3.2 Stimuli and design

A 2 [*Reporting Style*: direct versus indirect speech] × 2 [*Context*: loud-speaking versus quiet-speaking reported speaker] × 2 [*Word*: loud-speaking versus quiet-speaking spoken word] design was used in this experiment (see the example below; next page).

Twenty-four pairs of short fictitious stories with different protagonists were prepared as reading materials (they were exactly the same as those in Experiment 7). Each story contained either a direct speech or an indirect speech quotation (*Reporting Style* factor). As for each pair, one story described a situation where the protagonist would speak loudly whereas the other a situation where the speaker would speak quietly (*Context* factor). Crucially, the critical sentence (bold texts in the following example) within each quotation was identical across *Contexts* and virtually the same across *Reporting Styles*. Recordings of spoken words were taken from a female native English speaker and prepared as the probe stimuli. As for the 24 critical items, the spoken words matched the final words in the speech quotations (e.g., the spoken word '*neighbours'* matches the last word in *He looked very serious, demanding: 'Excuse me Madam. … from your* <u>neighbours</u>*.'*). They were acoustically normalised and manipulated into 'loud' and 'quiet' versions such that they either matched or mismatched the reported speakers' loudness which was contextually described in the stories (*Word* factor). Since the two versions of each spoken word were derived from the same recording file, they matched in all acoustic characteristics and only differed in loudness (i.e., vocal volume in the current investigation). The latter was confirmed by acoustic analysis on intensity: The loud-speaking words (M = 73.1 dB, SD =

9.6) were significantly louder than the quiet-speaking words (M = 56.6 dB, SD = 9.6), paired-sample $t(23) = 870.9$, $p < .001$.

		Loud-speaking word	Quiet-speaking word
Loud-speaking context	**Direct speech**	*Katy's friends were having such a great time blasting cheesy Christmas music in her flat that she barely heard the knock at the door at 1am. She was expecting another friend and was surprised to find a policeman looking crossly over her shoulder at the party going on inside. He looked very serious, demanding: 'Excuse me Madam. I suggest **you keep the noise down as there's been a complaint from your neighbours.**'*	*Katy's friends were having such a great time blasting cheesy Christmas music in her flat that she barely heard the knock at the door at 1am. She was expecting another friend and was surprised to find a policeman looking crossly over her shoulder at the party going on inside. He looked very serious, demanding: 'Excuse me Madam. I suggest **you keep the noise down as there's been a complaint from your neighbours.**'*
		Loud spoken probe word (e.g., 'neighbours')	**Quiet** spoken probe word (e.g., 'neighbours')
	Indirect speech	*Katy's friends were having such a great time blasting cheesy Christmas music in her flat that she barely heard the knock at the door at 1am. She was expecting another friend and was surprised to find a policeman looking crossly over her shoulder at the party going on inside. He looked very serious, demanding that **she should keep the noise down as there had been a complaint from her neighbours.***	*Katy's friends were having such a great time blasting cheesy Christmas music in her flat that she barely heard the knock at the door at 1am. She was expecting another friend and was surprised to find a policeman looking crossly over her shoulder at the party going on inside. He looked very serious, demanding that **she should keep the noise down as there had been a complaint from her neighbours.***
		Loud spoken probe word (e.g., 'neighbours')	**Quiet** spoken probe word (e.g., 'neighbours')
Quiet-speaking context	**Direct speech**	*Katy was reading in the library and had put ear plugs in to block out the other students studying nearby. She absentmindedly began to click her tongue, not realising that other people could hear her until a concerned librarian tapped her lightly on the shoulder. He politely said: 'Excuse me Madam. I suggest **you keep the noise down as there's been a complaint from your neighbours.'***	*Katy was reading in the library and had put ear plugs in to block out the other students studying nearby. She absentmindedly began to click her tongue, not realising that other people could hear her until a concerned librarian tapped her lightly on the shoulder. He politely said: 'Excuse me Madam. I suggest **you keep the noise down as there's been a complaint from your neighbours.'***
		Loud spoken probe word (e.g., 'neighbours')	**Quiet** spoken probe word (e.g., 'neighbours')
	Indirect speech	*Katy was reading in the library and had put ear plugs in to block out the other students studying nearby. She absentmindedly began to click her tongue, not realising that other people could hear her until a concerned librarian tapped her lightly on the shoulder. He politely said to her that **she should keep the noise down as there had been a complaint from her neighbours.***	*Katy was reading in the library and had put ear plugs in to block out the other students studying nearby. She absentmindedly began to click her tongue, not realising that other people could hear her until a concerned librarian tapped her lightly on the shoulder. He politely said to her that **she should keep the noise down as there had been a complaint from her neighbours.***
		Loud spoken probe word (e.g., 'neighbours')	**Quiet** spoken probe word (e.g., 'neighbours')

Moreover, 36 filler stories containing no direct or indirect speech were prepared and mixed with the 24 critical items to conceal the experimental manipulations of the study. The spoken words for the 36 filler items either matched or mismatched the selected critical words in the stories (not always the last words) and they were also presented in either a 'loud' or a 'quiet' version. Content-related comprehension questions were made for all items to encourage reading for comprehension. All reading materials except comprehension questions were segmented for serial visual presentation.

Eight lists of stimuli were generated with counterbalanced item-condition combinations (12 direct speech and 12 indirect speech items per list; 6 loudness match (3 loud, 3 quiet) and 6 loudness mismatch items (3 loud, 3 quiet) in each reporting style condition). Each critical item appeared only once per list, but in a different condition across lists. The eight lists of stimuli were evenly allocated to the participants. The presentation order of the stories on each list was randomised for each participant.

5.3.3 Procedure

The procedure was the same as in Experiment 6. Participants were seated comfortably in front of a Dell Dual Core PC, wearing headphones for spoken probe word presentation. Each participant was assigned to one of the eight lists containing different versions of experimental items, randomly interspersed with the filler stories. Stimulus presentation was implemented in E-prime v2.0 software (Psychology Software Tools, Inc., USA). Reading materials were presented in an 18 pt 'Arial Narrow' font printed in black on a light grey background. Each segment of text was presented at the centre of the screen.

Each trial began with the presentation of 'NEW TRIAL' in the middle of the screen for 1000 ms. It was followed by serial visual presentation of segmented text passages for silent reading. The presentation duration of each text segment was determined by 100 ms + 50 ms × the total number of syllables + 200 ms × the total number of words; this allowed sufficient time for reading. The text presentation would pause on a critical word while a spoken probe word was simultaneously presented. Participants had to respond as quickly and accurately as possible, using the mouse buttons (the left mouse button for 'yes', the right for 'no'), as to whether the word that they had read and the word that they had heard were the same or not. Their response times and accuracy were recorded. After reading,

participants had to answer a yes/no comprehension question regarding the content of the story they had read via pressing the corresponding mouse buttons. Answering the question triggered the presentation of the next trial.

Before the actual experiment, each participant had a 10-trial practice. In the actual experiment, they were allowed to have a short break after 30 trials.

5.3.4 Results and discussion

Incorrect responses to the spoken words were discarded (2.1%). Outliers were then excluded in two steps. First, the correct responses with a RT of longer than 2000 ms or shorter than 250 ms were removed. Next, the remaining trials with a RT which was 3SD away from a condition's mean were excluded. In total, 49 (3.3%) of the 1464 trials were excluded from analysis. The remaining 1415 trials were submitted to 2 [*Reporting Style*: direct versus indirect speech] × 2 [*Context*: loud-speaking versus quiet-speaking reported speaker] × 2 [*Word*: loud versus quiet versions of the probe words] ANOVAs with subject (F_1) and item (F_2) as random factors, respectively.

We found that the critical three-way interaction *Reporting Style* × *Context* × *Word* was not significant, Fs < 2.9, ps > .09. It suggests that loudness (i.e., vocal volumes) may not be encoded as part of the simulated 'inner voice' experiences during silent reading of direct as opposed to indirect speech, and that vocal loudness may be simulated in other types of representations such as prosody. No other effect was significant except that the *Reporting Style* × *Context* interaction was significant by subject, $F_1(1, 47) = 14.696$, $p < .001$, and marginally significant by item, $F_2(1, 23) = 4.030$, $p = .057$. Exploring this interaction, we found that the *Context* effects were not significant in the direct speech conditions whereas during silent reading of indirect speech, the participants responded significantly faster when reading the loud-speaking context than when reading the quiet-speaking context (see Table 8). The significant between-context difference in the indirect speech condition may reflect a 'baseline' difference in emotional intensity across *Context*: The loud-speaking stories are usually more intense and emotionally engaging (e.g., *shouting*) than the quiet-speaking stories (e.g., *speaking politely*); hence participants tended to respond faster while reading the former because they were more roused. However, other latent factors significantly diminished this 'emotional intensity' RT difference in the direct speech

conditions, which resulted in the observed *Reporting Style* × *Context* interaction. Such result patterns cannot be simply interpreted in mental simulations of 'supra-modal' emotional representations. This is because that direct speech is represented in a more vivid fashion than indirect speech; one would predict that individuals should be more likely (or no less likely) to mentally simulate emotion representations during silent reading of direct speech as opposed to indirect speech. If participants were more roused when reading loud-speaking indirect speech quotations than quite-speaking indirect speech quotations (as suggested by the significant *Context* effect in the indirect speech condition), they should have been even more roused or equally more roused during silent reading of loud-speaking direct speech quotations as compared to quiet-speaking direct speech quotations. However, the latter was clearly not the case: The RTs in the loud-speaking direct speech condition was significantly slower than the loud-speaking indirect speech condition (Table 8), and the *Context* contrast in the direct speech condition was not significant. Thus, latent factor which could exert inhibitory effects on RTs when reading loud-speaking direct speech quotations is needed to interpret the non-significant *Context* contrast in the direct speech condition. Suggested by our previous findings, one candidate could be mental simulations of voices. It has been evident that individuals mentally simulate voice representations during silent reading of direct speech but not indirect speech. The simulated voice representations are very likely to be acoustically distinct from the voices of the spoken probe words, even when they are congruent in loudness. This means that when our participants hear spoken probe words during silent reading of direct speech, they had to pay a cost to integrate the mentally simulated voices and the spoken words' voices before making a judgement response. Such a 'voice integration' cost was consistent with the significant RT delay in the loud-speaking, direct speech condition as compared to the loud-speaking, indirect speech condition. However, such a RT delay was not so significant in the quiet-speaking conditions. It perhaps suggests the reported speaker's loudness is represented in the salience of the simulated voice representations. If the 'quite' voice representations are less salient than the 'loud' voice representations, so would be the discrepancy between the simulated 'quiet' voices and the spoken words' voices. Thus, participants would pay less cost to integrate the two voice representations when reading quiet-speaking direct speech and this could explain why their RTs were not as much delayed as compared to when reading loud-speaking direct speech. Taken together, the *Reporting Style* × *Context* interaction could be interpreted as the combination of the

following effects (also see Figure 13 for an illustration): (1) The *Context* effects (i.e., faster RTs for the loud-speaking context) in the indirect speech condition reflected a between-context difference in emotional intensity; (2) such emotional intensity effects were diminished in the direct speech condition due to a more salient 'voice integration' cost in the loud-speaking context than in the quiet-speaking context.

			Direct Speech	Indirect Speech	Reporting Style Contrast			
					By subject		By item	
					t(60)	*p*	*t*(23)	*p*
Loudly-speaking Context			732 (203)	697 (176)	**3.94**	**< .001**	**2.31**	**.03**
Quietly-speaking Context			724 (202)	732 (203)	-1.15	.255	-.6	.55
Context Contrast	**By subject**	***t*(60)**	.77	**-3.45**	Note: Significant contrasts are highlighted in bold text			
		p	.44	**.001**				
	By item	***t*(23)**	.51	**-2.77**				
		p	.61	**.01**				

Table 8 RT means (in ms, alongside with standard deviations in parentheses) on the critical direct speech and indirect speech quotations across *Context*. Also shown are results of paired-sample *t*-tests investigating the simple contrasts of *Reporting Style* and *Context*.

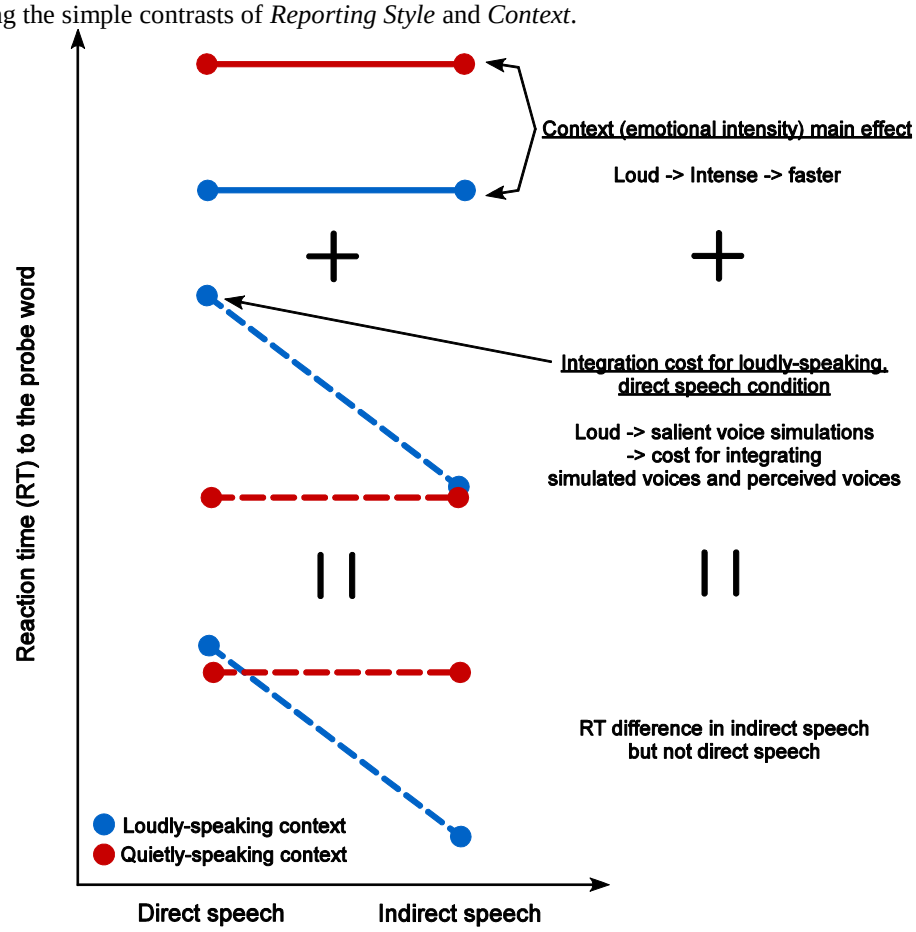


Figure 13 A potential interpretation of the *Reporting Style* × *Context* interaction in Experiment 8. The circles represent hypothetical RT means for the four conditions. **The top panel** illustrates the hypothetical result pattern when only the *Context* (emotional intensity) main effect is considered. **The middle panel** illustrates the hypothetical result pattern when only the voice simulation effect (the *Reporting Style* × *Context* interaction) is considered. **The bottom panel** illustrates the hypothetical result pattern when both the *Context* main effect and the voice simulation effects are considered – it resembles the results we found in Experiment 8.

In summary, the current data are in line with voice simulations during silent reading of direct speech but not indirect speech. They suggest that the reported speaker's loudness may not be encoded in the volume of his/her voice. Instead, it may be represented in terms of the *salience* of the simulated voice representations. Future research is needed to validate this 'salience' hypothesis of loudness representations. For example, we could measure the amplitude and the distribution of the brain activity within voice-selective areas in relation to loud- versus quiet-speaking contexts during silent reading of direct speech as opposed to indirect speech.

5.4 General discussion

The current chapter investigated the loudness aspect of the simulated 'inner voice' representations during oral reading (Experiment 7) and silent reading (Experiment 8) of direct versus indirect speech.

Experiment 7 showed that individuals spontaneously adjusted their own loudness in accordance with the contextually implied loudness of the reported speaker during oral reading of direct as opposed to indirect speech. In line with our hypothesis, the results could mean that individuals are more likely to mentally simulate the reported speaker's loudness during oral reading of direct speech as opposed to indirect speech.

However, mental simulations of loudness were not straightforwardly evident in silent reading. In Experiment 8, participants silently read story-embedded direct and indirect speech and responded to spoken words the volume of which either matched or mismatched the reported speaker's loudness that was linguistically described in the contexts. Inconsistent with our prediction, the participants' RTs were not effectively modulated either by the volume (intensity level) of the spoken words or by their congruency with the contexts. This could suggest that the paradigm was not sensitive enough to detect the simulated loudness representations because the probe word was presented when reading the final word of the speech sentence where voice simulations may have weakened over time. Future research should use a probe word that occurs early during reading of the speech utterance as opposed to at the final word of the sentence. Alternatively, it could suggest that the reported speakers' loudness may not be encoded in terms of the absolute vocal volume. Such results are consistent with Pitt and Crowder's (1992) research on

auditory imagery of tone loudness. In their investigations, they examined the effects of the subjectively-imagined loudness (instructed with written words *'loud'* or *'soft'*) of a sine wave tone on the subsequent pitch judgement on a probe tone which is either loudness-congruent or loudness-incongruent. They did not observe any loudness-congruency effect, either in RTs or in accuracy rates, suggesting that acoustic representations of loudness (volume) may not be encoded in loudness images. The authors hence proposed that in imagery, loudness may be encoded in another form (e.g., the prosodic information and the articulatory motoric information). In line with this speculation, we speculated that the simulated loudness in silent reading of direct versus indirect speech is also encoded in other forms. One such form, suggested by the significant *Reporting Style* × *Context* interaction observed in our silent reading investigation, may be the *salience* of the simulated voice representations. Another possibility, suggested by the contextual modulation of loudness we observed during *oral reading* of direct rather than indirect speech (Experiment 7), is that loudness may be encoded in mental simulations of loudness-related articulations, i.e., in a form of motor simulations. Nevertheless, further investigations are needed to verify how loudness is mentally represented during comprehension of direct versus indirect speech, using more sensitive methodologies such as fMRI.

Chapter 6 Contextual modulation of reading rates for direct versus indirect speech quotations

6.1 Introduction

The present chapter addresses another acoustic aspect (in addition to pitch and loudness) of the purported 'inner voice' experience during reading of quotations, namely the contextually implied *rate* of the quoted speech. Specifically, we examined whether readers adjust their reading rates in accordance with the contextually implied speech rate of the quoted speaker during both oral reading (Experiment 9) and silent reading (Experiment 10) of direct versus indirect speech. Moreover, by focusing on potential behavioural consequences of voice simulations during reading, the present studies represent an important extension of the brain-imaging research reported in Chapters 2 and 3.

It has been demonstrated that reading rates may reflect auditory imagery of speaking rates (Alexander & Nygaard, 2008). In their study, participants were exposed to pre-recorded voices of either fast or slow speakers. In subsequent reading trials, participants were prompted to imagine those speakers as authors of the given reading materials. It was found that both oral and silent reading rates were faster when participants imagined previously heard *fast speakers* as authors.

Although related, the present experiments are somewhat different. First, Alexander and Nygaard (2008) did not investigate any influences of reporting style (direct versus indirect speech), which are of paramount interest in our investigations. Second, their experimental setup may have encouraged auditory imagery during reading, focusing on how the auditory memory of a specific talker's speaking rate modulates reading rates. In contrast, our experiments will examine how *linguistic context* (implying either a fast-speaking or a slow-speaking quoted protagonist) modulates participants' reading rates during oral and silent reading of quotations. Just as in Yao et al. (in press), participants in our experiments will neither be exposed to auditory samples of specific speakers before reading, nor be instructed or encouraged to 'act out' or imagine voices. Moreover, the reading materials will only contain fictitious, unfamiliar speakers. Thus, our experiments are designed to tap into more *automatic* voice simulation processes in response to written quotations. Inspired by Alexander and Nygaard (2008), we assume that such automatic voice simulations

should be measurable in modulations of oral and silent reading rates, and informed by Chapter 2, we predict that these reading rate adjustments should be stronger for direct speech quotations than for meaning-equivalent indirect speech quotations.

In the following experiments, participants either orally (Experiment 9) or silently (Experiment 10) read short fictitious stories that contained either direct or indirect speech quotations. The quotations were embedded in linguistic contexts describing either a fast- or a slow-speaking reported speaker. Participants' oral and silent reading rates were measured via audio recordings and eye-tracking, respectively.

6.2 Experiment 9

6.2.1 Participants

Twenty native English speakers with no reported reading impairments participated in exchange for £2. A typical session lasted about 20 minutes.

6.2.2 Stimuli and design

A 2 [*Reporting Style*: direct versus indirect speech] × 2 [*Speaking rate*: fast- versus slow-speaking reported speaker] design was used (see an example below). Twenty-four quadruples of short fictitious stories were prepared as reading materials (see Appendix 4 for a complete stimulus set). The stories contained either a direct speech or an indirect speech quotation from a fictitious main protagonist in the story (*Reporting Style*). Across items, different protagonist names were used. A second variable (*Context*) was independently manipulated such that the quotations were preceded by a context that implied either a fast-speaking or a slow-speaking quoted protagonist. Crucially, critical quotation passages within each item (bold texts) were identical between the fast- and slow-speaking *Context* conditions, and nearly the same between the direct and indirect speech *Reporting Style* conditions. The quotations were always followed by at least one additional sentence.

	Fast-speaking context	Slow-speaking context
Direct speech	*It was a typical British day, rainy and gloomy. Sixteen year-old pianist Bobby was going to play in the quarter-finals of a local talent competition. He was extremely nervous before his performance. His mother encouraged him but he was all shaking and said: 'No! I can't do it! **This is the end of the journey because it is unlikely that I will make it this time**.' His mother tried to calm him down, saying that it's not the winning that counts, but the taking part.*	*It was a typical British day, rainy and gloomy. At Glasgow Royal Infirmary, an old man was dying, and too weak to sit up. His family members were sitting around the bed, feeling sad. He wanted to say something, so his daughter placed a cushion under his head. Slowly, he looked around and said: 'I'm grateful you're all here. **This is the end of the journey because it is unlikely that I will make it this time**.' Then he closed his eyes and everyone burst into tears.*
Indirect speech	*It was a typical British day, rainy and gloomy. Sixteen year-old pianist Bobby was going to play in the quarter-finals of a local talent competition. He was extremely nervous before his performance. His mother encouraged him but he was all shaking and said that he couldn't do it and that **it was the end of the journey because it was unlikely that he would make it this time**. His mother tried to calm him down, saying that it's not the winning that counts, but the taking part.*	*It was a typical British day, rainy and gloomy. At Glasgow Royal Infirmary, an old man was dying, and too weak to sit up. His family members were sitting around the bed, feeling sad. He wanted to say something, so his daughter placed a cushion under his head. Slowly, he looked around and said that he was grateful for their coming and that **it was the end of the journey because it was unlikely that he would make it this time**. Then he closed his eyes and everyone burst into tears.*

Four stimulus-lists with counterbalanced item-condition combinations (6 stories per condition per list) were constructed and word-processed into four versions of written scripts. Each item appeared in one condition per list, but in different conditions across lists. The order of the stories per script was randomised. Each script was randomly assigned to five participants.

6.2.3 Procedure

The experiment was conducted in a sound attenuated room. Participants were instructed to read out the stories from the script in one go and as naturally and fluently as possible. The instructions did not explicitly encourage participants to vocally 'act out' the stories during reading.

After setting up the microphone and digital voice recorder, the experimenter sealed the room and the participant started reading the stories aloud while audio recordings were taken. These were saved as wave-files for later analyses.

6.2.4 Results and discussion

All recordings were listened to individually by the experimenter. Around 5% of the trials were excluded from analysis because of word repetition or substitution errors. The remaining valid recordings were digitally visualized (in millisecond resolution) using GoldWave 4.26. For each trial, the temporal onsets and offsets of the critical quotation passages were identified. Next, oral reading rates (in numbers of syllables per second) for the critical quotations were calculated and submitted to 2 [*Reporting Style*: direct speech versus indirect speech] × 2 [*Context*: fast versus slow] ANOVAs with subject (F_1) and item (F_2) as random factors.

Apart from a *Context* main effect by subjects (F_1 (1, 19) = 4.74, $p < .05$, F_2 (1, 23) = 1.68, $p > .2$) there was a significant *Context* × *Reporting Style* interaction (F_1 (1, 19) = 8.90, $p < .01$, F_2 (1, 23) = 6.40, $p < .02$). Exploring the latter, it was found that direct speech quotations were read out significantly faster when the context implied a fast-speaking rather than a slow-speaking quoted protagonist. By contrast, no such contextual modulation of oral reading rate was found for indirect speech quotations (Table 9). This suggests that participants engaged in spontaneous vocal re-enactments of the reported speech act when reading aloud direct rather than indirect speech quotations, by adjusting their oral reading rates to the speech rates that were contextually implied. Importantly, the results cannot be explained by differences in length or morphological features (tense, person, etc.) between the two *Reporting Style* conditions because the interaction involved differential effects of *Context* within each *Reporting Style* condition.

			Direct Speech	Indirect Speech	Quoting Style Contrast			
					By Subjects		By Items	
					t(19)	*P*	*t*(23)	*p*
Fast-speaking Context			6.17 (1.0)	5.85 (0.9)	**3.02**	**0.007**	**2.41**	**.03**
Slow-speaking Context			5.88 (0.9)	5.93 (1.0)	-0.37	.71	-0.46	.65
Context Contrast	**By Subjects**	***t*(19)**	**3.65**	-0.37	Note: Significant contrasts are highlighted in bold			
		p	**.002**	.71				
	By Items	***t*(23)**	**2.77**	-1.13				
		p	**.01**	.27				

Table 9 Oral reading rates (in syllables per second, with standard deviations in parentheses) for the critical quotations in Experiment 9. Also shown are results from 2-tailed paired sample *t*-tests examining the simple effects of *Context* and *Reporting Style*.

6.3 Experiment 10

Experiment 10 was an eye-tracking study using the same experimental items as Experiment 9. The question was whether the results from Experiment 9 would replicate if participants were reading the written materials silently for comprehension.

6.3.1 Participants

Fifty-two native English speakers with no reported reading impairments participated in exchange for £3. A typical session lasted 30 minutes. Four participants were excluded from analysis due to poor eye-tracking (3 participants) or less than 75% answering accuracy (1 participant), leaving 48 participants for analysis.

6.3.2 Stimuli and design

The stimuli and design were identical to Experiment 9. The critical stories (24 items) were mixed with 24 filler stories, which contained no direct or indirect speech quotations to conceal the purpose of the study. Experimental and filler items spanned 4-7 lines on screen. Although it was impossible to keep the formatting absolutely identical, care was taken to ensure that the critical quotations spanned roughly the same screen areas across all four conditions within each item. Content-related questions were also prepared to encourage reading for comprehension.

6.3.3 Procedure

The experiment was conducted using a SR-Research EyeLink 1000 desk-mounted eye-tracker running at 1000 Hz sampling rate. Stimulus presentation was controlled by software developed at the University of Massachusetts (EyeTrack 0.7.1; www.psych.umass.edu/eyelab/software/). Participants were seated about 70 cm from a 21 inch CRT display running at 100 Hz refresh rate in 1280 × 960 pixel resolution. 3 characters equalled about one degree of visual angle. Materials were presented in an 18 pt Bitstream Vera Sans Mono font printed in black on a light grey background. Line spacing was set to 60 pt such that fixation locations could unambiguously be mapped onto a corresponding line of text. Viewing was binocular, but only the participant's dominant eye was tracked (as determined by a simple parallax test). A chin rest was used to keep

viewing distance constant and to prevent strong head movements during reading. Button responses were collected using a hand-held Microsoft USB game pad.

Each participant was assigned to one of four lists containing different versions of experimental items, pseudo-randomly interspersed with the filler stories. There were always two fillers at the beginning. At the start of the experiment, the standard EyeLink calibration and validation procedure was performed in which participants had to look at nine fixation targets in random succession. Calibration and validation were repeated during the experiment if the experimenter noticed a decline in measurement accuracy.

Each trial started with the presentation of a central fixation dot for drift-correction, followed by a small rectangle in the same location as the first character of the upcoming text display. A fixation for at least 200 ms on this rectangle triggered the presentation of the text, so that reading always started in the first character position. Participants were instructed to read the stories carefully and to press a button when they had finished reading. A simple yes/no comprehension question was then presented, which participants had to answer using either the left (‘yes') or the right (‘no’) response button. Answering the question trigged the presentation of the next trial.

6.3.4 Results and discussion

Fixation coordinates were mapped onto character positions using EyeDoctor 0.6.5 (www.psych.umass.edu/eyelab/software/). Fixations below 80 ms were pooled with temporally adjacent fixations if the latter were within half a degree of visual angle. Fixations on the critical quotation regions were summarized in terms of (i) *go-past time* (also known as *regression path duration*), i.e., the time from fixating the critical region for the first time until a subsequent region is fixated, and (ii) *no-regressions go-past time* (the same as before, but excluding trials where readers regressed back to an earlier region after initial reading of the critical region).

Measures \ Condition	Fast-speaking Direct	Fast-speaking Indirect	Slow-speaking Direct	Slow-speaking Indirect
Go-past reading time (ms)	1919 (805)	2099 (904)	2013 (853)	2059 (1001)
Number of go-past fixations	9.9 (3.9)	10.6 (4.2)	10.2 (4.0)	10.4 (4.4)
Go-past fixation duration (ms)	194 (34)	198 (31)	198 (32)	197 (31)

Table 10 Standard eye movement measures of processing across 4 conditions (with standard deviations in brackets).

The standard eye movement measures of processing (i.e., without removing 'outlier trials') across all 4 conditions were reported in Table 10. Outlier trials were removed in two steps. First, we excluded trials where the initial reading of the critical region was not fluent, operationally defined by a *Fixation Density Disparity* greater than two (see Appendix 5). Next, trials with a *go-past time* of more than 3 SD above an item's condition mean were removed. In total, 74 trials (6.4%, equally distributed across conditions) were excluded as a result of these two outlier definitions. Data from the remaining 1078 trials were submitted to 2 × 2 ANOVAs by participants and items.

There was a significant *Context × Reporting Style* interaction in *go-past time* ($F_1(1, 47) = 22.79$, $p < .001$, $F_2(1, 23) = 6.41$, $p < .02$): Direct speech quotations were read significantly faster when the context implied a fast-speaking rather than a slow-speaking protagonist, while no such *Context* contrast was found for indirect speech quotations (Table 11). In *no-regressions go-past time* (excluding 16% of trials where readers regressed back to an earlier region) the same *Context × Reporting Style* interaction emerged ($F_1(1, 47) = 14.90$, $p < .001$, $F_2(1, 23) = 4.91$, $p < .04$, Table 12). The latter confirms that the *go-past time* effects in Table 11 were not driven by re-inspections of earlier regions.

The by-subject *t*-tests in Table 11 and Table 12 also suggest a 'reversed' (fast > slow) simple effect of *Context* in the indirect speech condition. Closer inspection revealed that this contrast was due to only three stimuli (items 1, 23, and 24 in Appendix 4). After removing those stimuli from analysis, the *Context × Quoting Style* interaction remained significant (*go-past time*: $F_1(1, 47) = 7.50$, $p < .01$, $F_2(1, 20) = 4.65$, $p < .05$; *no-regressions go-past time*: $F_1(1, 47) = 6.00$, $p < .01$, $F_2(1, 20) = 3.90$, $p = .062$), but crucially, the *Context* effect in the indirect speech condition disappeared (Table 13 and Table 14). It is therefore hard to interpret the latter as anything other than item-specific

idiosyncrasy. By contrast, the *Context*-effect in the direct speech condition (fast < slow) robustly showed up in all analyses.

In conclusion, not only oral readers (Experiment 9), but also silent readers adjust their reading rates to the contextually implied speech rate when reading direct speech as opposed to indirect speech quotations. Again, the observed interactions cannot be explained by differences in length and/or morphology between the two *Reporting Style* conditions. Additional analyses of numbers and average durations of go-past fixations (Appendix 6) revealed that the reported *go-past time* effects were mainly carried by fewer fixations being made in the fast-speaking conditions. Average durations of go-past fixations did not seem to be affected much by the experimental manipulations.

			Direct Speech	Indirect Speech	Quoting Style Contrast			
					By Subjects		By Items	
					t(47)	*p*	*t*(23)	*p*
Fast-speaking Context			1857 (743)	2056 (885)	**-3.17**	**.003**	**-2.56**	**.02**
Slow-speaking Context			2011 (830)	1942 (896)	0.50	.62	0.29	.78
Context Contrast	**By Subjects**	***t*(47)**	**-3.50**	1.84	Note: Significant contrasts are highlighted in bold			
		p	**.001**	.07				
	By Items	***t*(23)**	**-2.32**	1.43				
		p	**.03**	.17				

Table 11 *Go-past* times (in ms, with standard deviations in brackets) for the critical quotations in Experiment 10 (silent reading). Also shown are results from 2-tailed paired sample *t*-tests examining the simple effects of *Context* and *Reporting Style.*

			Direct Speech	Indirect Speech	Quoting Style Contrast			
					By Subjects		By Items	
					t(47)	*p*	*t*(23)	*p*
Fast-speaking Context			1784 (693)	1957 (853)	**-3.42**	**.001**	**-2.43**	**.02**
Slow-speaking Context			1932 (799)	1845 (820)	1.01	.32	0.04	.97
Context Contrast	**By Subjects**	***t*(47)**	**-3.06**	**2.55**	Note: Significant contrasts are highlighted in bold			
		p	**.004**	**.01**				
	By Items	***t*(23)**	**-2.42**	1.16				
		p	**.02**	.26				

Table 12 *No-regressions go-past* times (in ms, with standard deviations in brackets) for the critical quotations in Experiment 10 (silent reading). Also shown are results from 2-tailed paired sample *t*-tests examining the simple effects of *Context* and *Reporting Style*.

			Direct Speech	Indirect Speech	Quoting Style Contrast			
					By Subjects		By Items	
					t(47)	*p*	*t*(20)	*p*
Fast-speaking Context			1919 (752)	2043 (897)	-1.90	.06	**-2.13**	**.05**
Slow-speaking Context			2081 (828)	2024 (915)	0.35	.73	0.63	.54
Context Contrast	**By Subjects**	***t*(47)**	**-3.81**	0.25	Note: Significant contrasts are highlighted in bold			
		p	**.001**	.80				
	By Items	***t*(20)**	**-2.78**	0.37				
		p	**.01**	.72				

Table 13 ***Go-past times*** **(in ms, with standard deviations in brackets) for the critical quotations in Experiment 10 (silent reading), after removing items 1, 23, and 24 from analysis.** Also shown are results from 2-tailed paired sample *t*-tests examining the simple effects of *Context* and *Reporting Style*.

			Direct Speech	Indirect Speech	Quoting Style Contrast			
					By Subjects		By Items	
					t(47)	*p*	*t*(20)	*p*
Fast-speaking Context			1843 (701)	1941 (860)	**-2.14**	**.04**	-2.00	.06
Slow-speaking Context			1996 (800)	1904 (838)	1.45	.15	0.64	.53
Context Contrast	**By Subjects**	***t*(47)**	**-3.27**	0.85	Note: Significant contrasts are highlighted in bold			
		p	**.002**	.40				
	By Items	***t*(20)**	**-2.80**	0.33				
		p	**.01**	.75				

Table 14 ***No-regressions go-past*** **times (in ms, with standard deviations in brackets) for the critical quotations in Experiment 10 (silent reading), after removing items 1, 23, and 24 from analysis.** Also shown are results from 2-tailed paired sample *t*-tests examining the simple effects of *Context* and *Reporting Style*.

Finally, we also performed by-item correlation analyses to evaluate the consistency of the results between the two experiments. Indeed, the raw oral reading times (in ms) from Experiment 9 correlated well with the silent reading data from Experiment 10 ($r(94) = .86$, $p < .001$)[14]. Given that these correlations are partly due to length effects (longer quotations take longer to read both orally and silently), we also performed multiple regression analyses with silent reading times (Experiment 10) as criterion variables, and both length (in numbers of words) and raw oral reading times (Experiment 9) as simultaneous predictors. These analyses obtained *partial correlations* of $r(93) = .56$ ($p < .001$), which means that the two experiments obtained very consistent by-item results even when length effects were partialled out.

[14] We considered both *go-past* and *no-regressions go past times* in these correlation analyses. Rounded to the second decimal, the *r*-values were identical.

6.4 Summary and general discussion

In two experiments, we found that both oral and silent readers adjusted their reading rates in accordance with the contextually implied speech rate of a quoted protagonist, but only when quotations employed a direct rather than an indirect speech quotation style. In line with our previous research using fMRI (Chapter 2), the current findings (particularly from Experiment 10) support a hypothesised 'inner voice' experience during silent reading of direct speech quotations; more specifically, we showed that speech rate is an important aspect of this inner voice experience and that mental simulations of voice are not only reflected in brain-activation patterns and vocalisations but also in behavioural eye-movement patterns.

The present results also extended previous findings by Alexander and Nygaard (2008) who found that pre-exposure to the speech rates of specific speakers influences how fast oral and silent readers would process written stimuli that were supposedly authored by those speakers. Specifically, the present experiments did not focus on this kind of speaker adaptation. Instead, we were interested in how information provided by the linguistic context (implying either a fast-speaking or a slow-speaking quoted protagonist) influences participants' reading rates during oral and silent reading, and importantly, how different reporting styles would affect such contextual modulations of reading rate. Another aspect in which our studies differed from Alexander and Nygaard's (2008) – apart from using eye-tracking instead of self-paced reading for silent reading – is that our experimental instructions did not explicitly encourage participants to 'act out' or imagine speech during reading (see 6.1 Introduction). In Experiment 9 (oral reading) participants were just asked to read the stories aloud and in Experiment 10 (silent reading) experimental tasks emphasised reading for comprehension. The observed contextual modulations of reading rate in response to direct speech quotations therefore suggest that, in this type of quotations, readers *routinely* and *automatically* activate mental representations that link to how the quoted speaker would sound like (in this case, how fast the quoted protagonist would speak), which in turn have an effect on the rate of information uptake (silent reading) and delivery (oral reading). In this respect, the present results are an important extension of our previous findings (e.g., Chapter 2), who equally found that silent readers of direct speech quotations routinely and automatically activate voice-related perceptual representations.

Our findings may be interpreted in terms of *perceptual simulation*, as proposed by embodied cognition theories (Barsalou, 1999, 2008). Such theories argue that mental representations of language are grounded in perceptual experiences and actions, and that perceptual simulation (i.e., the mental re-enactment of perceptual, motor, and introspective states acquired during experience with the world, body and mind) is an automatic and integral part of language comprehension. Under such a premise, one could argue that accumulated experiences with how direct versus indirect speech quotations are typically *used* form the basis for differential degrees of perceptual simulation during language comprehension. When speakers employ direct speech, they often mimic or dramatize aspects of the reported speaker's voice in order to depict the reported speech act; indirect speech, by contrast, is typically not used in such a demonstrative fashion (Clark & Gerrig, 1990). Comprehension of direct speech is therefore more likely to be grounded in the perceptual experience of a vocal demonstration or dramatization of a reported speaker's utterance, and thus more likely to invoke perceptual simulations of the reported speaker's voice. Put differently, a direct-speech reporting style is more likely to be taken as a cue to spontaneously engage in vivid perceptual simulations of the reported speech act than an indirect-speech reporting style.

Of course, there are still a number of open questions which our research cannot answer at present. Some are related to the nature of voice simulation itself. For instance, is the level at which voice simulation operates primarily prosodic (especially given that loudness may not be encoded in such processes; cf. Chapter 5), and what kind of processing takes place in conditions where voice simulation is reduced or absent? Other questions relate to eye-movement control in reading. The results from Experiment 10 suggest that high-level factors such as context and quoting style have an impact on basic reading rates. Do these pragmatically induced reading-rate modulations operate at the level of individual word processing or at the level of more 'global' reading parameters? As these questions indicate, further research into this area might substantially deepen our understanding of written language processing as a whole.

In conclusion, our investigations address the pragmatics of different quotation styles and their associated mental representations during text comprehension. Our experiments showed that oral and silent readers spontaneously adjust their reading rates to contextually implied speech rates when reading direct speech as opposed to meaning-equivalent indirect

speech quotations. These reading-rate adjustments may be interpreted as perceptual simulations of reported speakers' voices. Specifically, it suggests that temporal representations such as speaking rate are an important aspect of such simulated 'inner voice' experience in reading direct speech as opposed to indirect speech.

PART 3: The Visual Aspect of the Direct-Speech Representations: Mental Simulations of Facial Expressions?

Chapter 7 Are facial expressions also encoded in silent reading of direct speech? (Experiment 11)

7.1 Introduction

Using fMRI, we have shown that individuals are more likely to mentally simulate the reported speakers' voices during language comprehension of direct as opposed to indirect speech (Chapter 2); the simulated voice representations entail supra-segmental acoustic information (Chapter 3). At the behavioural level, we found that the simulated voice representations could be characterised in modulations of acoustic parameters such as speaking rate. The findings are in line with Clark and Gerrig's (1990) demonstration theory of direct speech and the embodied theories in language processing (e.g., Barsalou, 1999, 2008; Zwaan, 2004).

However, in verbal communication, direct-speech demonstrations usually entail not only vivid depictions of the reported speakers' voices but also depictions of other non-vocal paralinguistic information such as the reported speaker's facial expressions. Hence, a thorough investigation of the mental representations during silent reading of direct versus indirect speech should, in addition to voices, demonstrate simulations of other non-voice aspects such as facial expressions. The literature on recognition of facial expressions suggests that understanding facial expressions involves embodied simulations of facial expressions by the perceivers themselves (see Niedenthal, Mermillod, Maringer, & Hess, 2010 for a review and the proposal of the Simulation of Smiles (SIMS) model). Such a 'simulation' mechanism of understanding facial expressions has recently been demonstrated in language comprehension (Havas, Glenberg, Gutowski, Lucarelli, & Davidson, 2010). The authors applied subcutaneous injections of botulinum toxin-A (BTX) to temporarily paralyse participants' facial muscles and asked them to read sentences (e.g., *The pushy telemarketer won't let you return to your dinner*) implying certain emotions (e.g., anger). They found that the BTX treatment selectively slowed down the reading of sentences that described situations which require the paralysed muscle for expression the emotions evoked by the sentences. Their results indicate that processing of emotional language requires embodied simulations of facial expressions and causally activates facial muscles that are used in expressing the relevant emotions. Regarding comprehension of direct versus indirect speech, re-analysis of our previous fMRI results (Chapter 2) revealed

that silent reading of direct speech elicited higher brain activations not only within the voice-selective areas of the auditory cortex but also within the occipital cortex, superior parietal lobules and precuneus (Figure 14A). Activation of these non-voice areas is in line with an enriched multi-sensory mental simulation process for direct speech that also encompasses, for instance, visual aspects of the described situation. When we applied a face localiser[15] on the whole-brain contrast, we observed that the brain areas that are more active when seeing face images as opposed to non-face images (e.g., the bilateral fusiform face areas (FFA); cf. Kanwisher, McDermott, & Chun, 1997, and the right posterior superior temporal sulcus (pSTS)) were also more active during silent reading of direct speech as opposed to indirect speech (Figure 14B). The findings indicate the possibility that individuals may indeed mentally simulate face representations during silent reading of direct speech but not indirect speech.

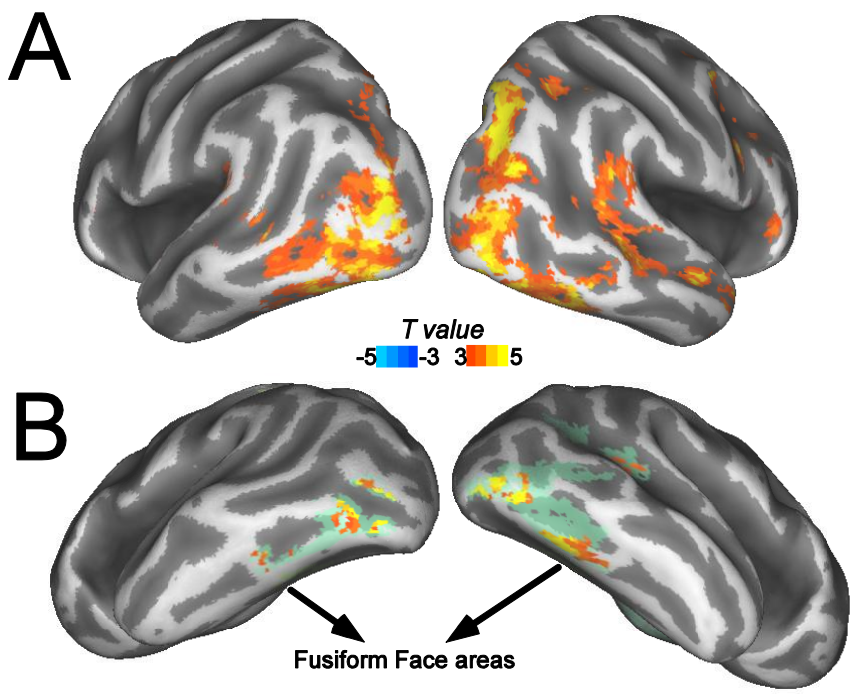


Figure 14 The brain activations in the contrast of direct speech as opposed to indirect speech in Experiment 1 (Chapter 2). A, the whole-brain contrast, **B,** the between-condition differences masked by a face localiser (the green areas indicate the brain areas that are more sensitive to seeing face images than to seeing non-face images). Only the group ROI is used because the subjects were from an independent sample.

[15] The face localiser was obtained from a different sample of 18 participants. During the localiser task, participants were instructed to look at images that are presented at the centre of the screen. The 5 20-sec blocks of faces, 5 20-sec blocks of face-like houses and 5 20-sec blocks of scrambled visual patterns were presented in an efficiency optimised, pseudo random order. The face-sensitive brain areas were obtained by contrasting the face condition to both the house condition and the visual pattern condition.

Thus, in this chapter, we ventured beyond the 'inner voice' experiences by investigating whether individuals are more likely to mentally simulate reported speakers' facial expressions during silent reading of direct as opposed to indirect speech. Since facial expressions are important vehicles of emotions in human interactions (Matsumoto, Keltner, Shiota, O'Sullivan, & Frank, 2008; Russell, Bachorowski, & Fernandez-Dols, 2003), we manipulated the types of facial expressions by manipulating the reported speakers' emotions. We examined three emotions, joy, surprise and anger, which are of similar arousal levels and are of either positive or negative valence; they would presumably cover the range of facial expressions in a relatively comprehensive way. We adopted the 'reading & recognition' paradigms in the literature on visual simulations (e.g., Connell, 2007; Stanfield & Zwaan, 2001; Zwaan et al., 2002). Specifically, we examined the effects of the linguistically implied facial expressions of the reported speakers on the recognition latencies to visual stimuli of facial expressions that were made relevant by the previous text passages. We prepared short stories containing direct or indirect speech in which the reported speakers were either happy, or surprised, or angry. And we prepared graphic emoticons which illustrate facial expressions of the investigated emotions as the visual probe stimuli to match the stories. We chose graphic emoticons over photorealistic face images as the visual probe stimuli because the former are highly abstract in the confounding identity-specific information such as the person's age, gender, skin type, etc. that we were not interested in, and yet they still illustrate sufficiently detailed facial expressions and have been found to activate face-sensitive brain areas such as the fusiform face areas (FFA) in a similar manner as the latter (Yuasa, Saito, & Mukawa, 2011). Participants silently read a series of short stories containing direct or indirect speech, each followed by a graphic emoticon, and answered 'yes' or 'no' to indicate whether the illustrated expression was congruent with the reported speaker's facial expression that was implied in the preceding text passage. If individuals mentally simulate reported speakers' facial expressions during silent reading of direct speech rather than indirect speech, their recognition on expression-congruent graphic emoticons should be facilitated whereas their recognition on expression-incongruent graphic emoticons should be inhibited, but only after reading direct speech not indirect speech.

In addition, we manipulated linguistic focus on reported speakers' facial expressions to verify the nature of the underlying representations of the purported differential face

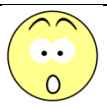
simulations during silent reading of direct versus indirect speech. Similar to the *Linguistic Focus* manipulation on voices in Experiment 3, we employed *face-focused* versus *face-unfocused* linguistic descriptions before the direct and indirect speech quotations (e.g., *Vicky smiled jubilantly: '...'* versus *Vicky felt rather relieved: '...'*). Advised by the results in Experiment 3, we predicted any observed differential face congruency effects during silent reading of direct versus indirect speech should be selectively masked by exerting linguistic focus on faces (see the masking hypothesis in 4.1 Introduction and see 4.2.5 for the evidence of the masking hypothesis on voice simulations). In other words, face-focused descriptions should activate face representations (e.g., smiling faces from *Vicky smiled jubilantly*) which would be integrated during silent reading of both direct and indirect speech. Thus, we should observe a face congruency main effect regardless of the reporting styles. By contrast, the purported differential face simulations in silent reading of direct versus indirect speech should be retained when the quotations are preceded with face-unfocused descriptions which do not activate face representations. And we should observe an interaction between the face congruency factor and the reporting style factor. In sum, if we observe that the differential face congruency effects during silent reading of direct versus indirect speech can indeed be selectively masked by exerting the linguistic focus on faces, we can conclude that the representations underlying such congruency effects should contain face representations.

7.2 Participants

Sixty-four native speakers of English participated in this experiment (21 male, 43 female, age 17 – 37, M = 21.6, SD = 3.5). They had no reported reading or relevant cognitive impairments. They were paid £3 for their participation. A typical session lasts about 25 min.

7.3 Stimuli and design

A 2 [*Reporting Style*: direct versus indirect speech] × 2 [*Face Match*: the reported speaker's facial expression *matches* versus *mismatches* the facial expression of the emoticon] × 2 [*Linguistic Focus*: *face-focused* versus *face-unfocused* pre-quotation descriptions] design was employed (see below for an example).

		Story	Probe Picture	
		In her dinner party, Vicky was a bit nervous because she had prepared some exotic recipes that she had never attempted before. Unexpectedly, all dishes went extremely popular and some guests even craved more after they finished.	**Face Match**	**Face Mismatch**
Face Focused	**Direct speech**	*Vicky smiled jubilantly: '**Wow, I didn't expect that. I am so glad that you all enjoyed it!**'*		
	Indirect speech	*Vicky smiled jubilantly, saying that **she did not expect that and she was so glad that they all enjoyed it**.*		
Face Unfocused	**Direct speech**	*Vicky felt rather relieved: '**Wow, I didn't expect that. I am so glad that you all enjoyed it!**'*		
	Indirect speech	*Vicky felt rather relieved, saying that **she did not expect that and she was so glad that they all enjoyed it**.*		

Thirty-six short stories with different protagonists were prepared as reading materials. The stories described protagonists who were either happy (12 stories), or angry (12 stories), or surprised (12 stories). Half of the 'surprised' stories described positive events (e.g., finding an antique) whereas the other half described negative events (e.g., receiving news of death). Together with the 'happy' and 'angry' stories, half of the 36 critical stories described reported speakers of positive emotions while the other half reported speakers of negative emotions. Each story started with two declarative sentences to set up a scenario, followed by either a direct speech or an indirect speech from a main protagonist in the story (*Reporting Style* factor). The quotation was preceded by either a face-focused description or a face-unfocused description (*Linguistic Focus* factor). Three graphic emoticons illustrating happy, surprised and angry emotions were created to accompany the stories (see Figure 15); they either matched or mismatched the reported speaker's facial expressions that were linguistically implied in the story (*Face Match* factor). For each emotion, there was equal number of two mismatch emoticons: For instance, half of the 'happy' emotion items had the 'surprised' emoticon as the mismatch stimulus whereas the other half of the items had the 'angry' emoticon as the mismatch stimulus. The probe pictures were sized 237 × 237 pixels (ca. 9 inches^2) and the graphic emoticons were of the same colour, size and stylistic style. Yes/no comprehension questions regarding the contents of the stories (e.g., *Was Vicky's dinner party successful?*) were also prepared to

assess participants' overall comprehension performance and to ensure that they read the stories attentively.

In addition to the 36 critical items, 42 filler stories were prepared to conceal the intended experimental manipulations. They were accompanied with pictures of *non-face* objects (e.g., a book) which were either congruent or incongruent with the corresponding stories. Nine of the filler stories contained direct speech quotations and nine contained indirect speech quotations. The remaining 24 filler items were declarative discourses with no reported speech. In total, there were 54 stories which contained direct or indirect speech; 36 of them were accompanied with graphic emoticons (critical items) while 18 of them accompanied with pictures of *non-face* objects. Yes/no comprehension questions were also prepared for filler items.

Eight lists of counterbalanced item-condition combinations were yielded. Each item appeared in one condition per list and in different conditions across lists. Each participant only saw one list. The presentation order was randomised for each participant.

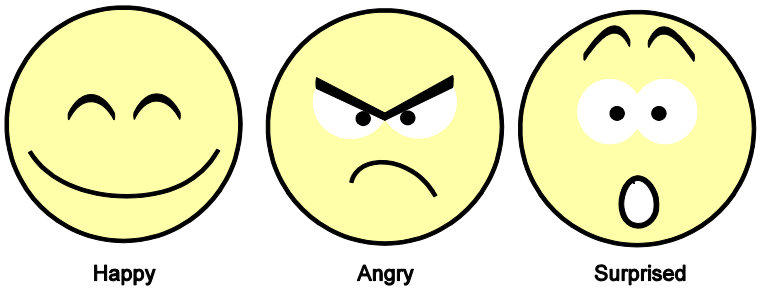


Figure 15 Graphic **emoticons illustrating 'happy', 'angry' and 'surprised' expressions.** They were of the same size, colour and style.

7.4 Procedure

Each participant was presented with one of the eight stimulus lists using E-prime 2.0 presentation software (Psychology Software Tools, Inc., USA). In each trial, the participant first saw a cue 'NEW TRIAL' in the middle of the computer screen for 1000 ms, which was followed by the presentation of a short story. The text was presented in an 18 pt Calibri font on a light grey background. The whole text passage was centred and wrapped within 65% width of the screen, together making the reading stimuli appear as natural as possible. Participants were instructed to silently read the text passage in their own time, and press spacebar to terminate the presentation as they finished reading. Next,

they saw a question mark at the centre of the screen for 1000 ms, followed by a graphic emoticon. They were instructed to indicate as quickly and accurately as possible whether or not the illustration (of either a facial expression or an object) was congruent with the story they had read, by pressing mouse buttons (left for 'yes', right for 'no'). At the end of each trial, participants were prompted to answer a yes/no comprehension question regarding the story they had just read.

7.5 Results and discussion

The participants' accuracy on the comprehension questions was all above 75% (M = 89%, SD = 31.4%), indicating that they read the stories attentively. Surprisingly, the participants' accuracy on the emoticon recognition amounted to a mere 87% (SD = 34%), which was even lower than the accuracy on the comprehension questions. This figure was unexpectedly low because previous investigations using similar 'reading & recognition' paradigms (e.g., Stanfield & Zwaan, 2001) usually report a recognition accuracy of higher than 95%. Thus, to ensure that our experimental manipulations were effective as intended, we needed to identify the source and the cause of the recognition errors before we proceed to test our hypothesis. To this end, we first conducted a logit binomial GEE (Hardin & Hilbe, 2003) on participants' recognition accuracy (as the binary dependent variable), with *Reporting Style*, *Context Emotion* (happy, surprised, and angry), *Emoticon* (happy, surprised and angry) and *Linguistic Focus* as predictors. There was a significant *Context Emotion* main effect, *Generalised Score Chi-Square*(2) = 25.48, $p < .001$, a significant *Emoticon* main effect, *Generalised Score Chi-Square*(2) = 21.89, $p < .001$, a significant *Context Emotion* × *Emoticon* interaction, *Generalised Score Chi-Square* (4) = 47.47, $p < .001$, and a significant *Context Emotion* × *Emoticon* × *Reporting Style* × *Linguistic Focus* interaction, *Generalised Score Chi-Square*(4) = 11.60, $p = .021$. The results indicate that the participants' recognition errors were unevenly distributed across the *Context Emotion* and the *Emoticon* manipulations. To illustrate these distributions, we summarised the recognition error rates by these two factors in Table 15, alongside with their corresponding RTs reported in Table 16. We found that the recognition errors were distinctly made in the 'surprised'-related conditions, irrespective of whether the emotion was conveyed by the linguistic context or by the emoticon. Consistently, the RTs in these 'surprised' conditions were also much longer than the 'surprise-free' conditions. Such error distribution patterns suggest that our participants had difficulty in processing the 'surprised' expression which

they tended to confused with the 'happy' and the 'angry' expressions. Such confusions indicate that the *Face Match* manipulations in the 'surprised'-related trials were not as effective as intended. For this reason, we had to exclude all 'surprised'-related trials in order to examine the effects of *Face Match*, which only left us 12 out of the 36 items for the following analyses.

The mean error rates (%)		**Context Emotion**			
		Happy	***Surprised***	**Angry**	**Mean**
Emoticon	**Happy**	1.04 (10.17)	***42.19 (49.51)***	3.65 (18.79)	11.98 (32.49)
	Surprised	***34.90 (47.79)***	***10.94 (31.25)***	***21.88 (41.45)***	19.66 (39.77)
	Angry	0 (0)	***22.92 (42.14)***	5.21 (22.25)	8.33 (27.66)
	Mean	9.24 (28.98)	21.74 (41.28)	8.98 (28.61)	

Table 15 Mean error rates (in %, along with standard deviations in parenthese) across the *Context Emotion* conditions and the *Emoticon* conditions. Unexpectedly high error rates (> 10%) were highlighted in bold text.

The RTs (ms)		**Context Emotion**			
		Happy	***Surprised***	**Angry**	**Mean**
Emoticon	**Happy**	671 (341)	***899 (701)***	800 (574)	755 (521)
	Surprised	***1096 (656)***	***966 (652)***	***1144 (686)***	1043 (665)
	Angry	770 (320)	***952 (527)***	863 (481)	862 (463)
	Mean	802 (470)	946 (636)	913 (578)	

Table 16 Mean RTs (in ms, along with standard deviations in parenthese) across the *Context Emotion* conditions and the *Emoticon* conditions. Also highlighted are the RTs in the conditions with high error rates (see Table 15).

Incorrect responses to the spoken words were discarded from the valid 12 items (2.5%). The outliers were excluded in two steps: First, the correct responses with a RT of longer than 3000 ms or shorter than 300 ms were removed. Next, the remaining trials with a RT which was 3SD away from an item's condition mean were excluded. In total, 4.8% (37 out of 768 trials) of the trials were excluded from analyses.

The remaining trials were submitted to 2 [*Reporting Style*: direct versus indirect speech] × 2 [*Linguistic Focus*: face-focused versus face-unfocused pre-quotation descriptions] × 2 [*Face Match*: the reported speaker's facial expression matches versus mismatches the facial expression in the emoticon][16] ANOVAs with subject (F_1) and item (F_2) as random factors, respectively. We found that the *Reporting Style* × *Face Match* interaction (*p*s > .26) and the *Linguistic Focus* × *Reporting Style* × *Face Match* interaction (*p*s > .9)

[16] We merged the *Context Emotion* factor and the *Emoticon* factor into the *Face Match* factor to simplify the interpretations of the effects.

were not significant. However, considering the inevitable reduction in statistical power after discarding two third of our data, we nonetheless explored the patterns of these interactions to assess whether they showed trends that were consistent with our hypothesis. As for the *Reporting Style* × *Face Match* interaction, we found that after reading direct speech, the participants responded 44 ms faster to the face-match emoticons as opposed to the face-mismatch emoticons whereas in the indirect speech conditions, the RT difference across the *Face Match* conditions was much less robust and in the opposite direction (the match condition was 7 ms slower than the mismatch condition; see Table 17). This result pattern was consistent with our hypothesis that individuals are more likely to mentally simulate the reported speakers' facial expressions in silent reading of direct as opposed to indirect speech. As for the *Linguistic Focus* × *Reporting Style* × *Face Match* interaction, we split the data by *Linguistic Focus* and conducted two separate *Reporting Style* × *Face Match* ANOVAs upon the face-focused data set and the face-unfocused data set, respectively. We found that when the speech quotations were preceded with face-unfocused descriptions, the *Reporting Style* × *Face Match* interaction was significant by item, $F_2(1, 11) = 6.89$, $p = .023$, and was approaching significance by subject, $F_1(1, 52) = 2.41$, $p = .127$. In the direct speech condition, RTs were significantly faster (by 97 ms; significant by subject and marginal by item) in the face-match condition than in the face-mismatch condition; whereas in the indirect speech condition, there was virtually no difference (amounting to a mere 1 ms) between the face-match and -mismatch conditions (see the upper panel in Table 18). As for the face-focused trials, such interaction was 'masked' (all *p*s > .6); the RTs in the face-match condition was numerically slower (by on average 12 ms) than in the face-mismatch condition, independent of the reporting style factor. Consistent with our predictions, the results indicate that the *Reporting Style* × *Face Match* interaction can be selectively masked by exerting linguistic focus on faces. This 'face-selective' masking effect suggests that this interaction reflects mental simulations of *facial expressions* (as opposed to *non-face* representations) during silent reading of direct rather than indirect speech. When reconciled with the face-related brain activations in silent reading of direct rather than indirect speech (Figure 14), the current findings (although some results were not significant due to the exclusion of 'surprised-related' trials) indicates the possibility that individuals may indeed mentally simulate facial expressions of the reported speakers during silent reading of direct rather than indirect speech.

Thus, future research would be needed to replicate the current result patterns, with three aspects of experimentation improved. First, given that the 'surprise' facial expressions are not so clearly defined (both in terms of emoticons and in terms of simulations), the experimental stimuli should perhaps only contain stories describing 'happy' and 'angry' emotions. Second, the current experimental setup required participants to press the left mouse button for "yes" and the right mouse button for "no". It is known that positive emotions are typically (for right-handers) associated with rightward space whereas negative emotions with leftward space (Daniel, 2009; Holmes & Lourenco, 2011). Hence in our future research, we will counterbalance the button settings (i.e., half of the participants press the left mouse button for "yes" while the other half press the right mouse for "yes") to prevent potential 'button-pressing' bias towards emotionally-valenced stimuli used in our investigations. Finally, we will also norm our reading materials in terms of semantic equivalence to ensure that the *Linguistic Focus* manipulation is not confounded with semantic discrepancies between the *Focus* and the *non-Focus* conditions.

			Direct Speech	Indirect Speech	Quoting Style Contrast			
					By Subjects		By Items	
					t(63)	*p*	*t*(11)	*p*
Face Match			691 (262)	741 (306)	-1.56	.12	-1.39	.19
Face Mismatch			735 (232)	734 (218)	.165	.87	.117	.91
Context Contrast	**By Subjects**	***t*(63)**	-1.66	.21	Note: Significant contrasts are highlighted in bold			
		p	.10	.84				
	By Items	***t*(11)**	-.798	.21				
		p	.44	.84				

Table 17 Mean RTs (in ms, along with standard deviations in parenthese) in the face match and face mismatch conditions across *Reporting Style*. Also shown are the paired-sample *t*-test results examining the simple effects of *Reporting Style* and *Face Match*.

Face-unfocused trials (e.g., *Vicky felt rather relieved*)			Direct Speech	Indirect Speech	Quoting Style Contrast			
					By Subjects		By Items	
					t(63)	*p*	*t*(11)	*p*
Face Match			651 (244)	715 (298)	-1.01	.32	-2.08	.06
Face Mismatch			748 (237)	716 (185)	**2.24**	**.03**	1.12	.29
Context Contrast	By Subjects	*t*(63)	**-2.51**	.42	Note: Significant contrasts are highlighted in bold			
		p	**.015**	.68				
	By Items	*t*(11)	-2.02	.046				
		p	.069	.96				

Face-focused trials (e.g., *Vicky smiled jubilantly*)			Direct Speech	Indirect Speech	Quoting Style Contrast			
					By Subjects		By Items	
					t(63)	*p*	*t*(11)	*p*
Face Match			732 (273)	766 (313)	-.78	.44	-.65	.53
Face Mismatch			722 (228)	752 (246)	-.13	.90	-.88	.40
Context Contrast	By Subjects	*t*(63)	-.04	.45	Note: Significant contrasts are highlighted in bold			
		p	.97	.66				
	By Items	*t*(11)	.19	.29				
		p	.85	.78				

Table 18 Mean RTs (in ms, along with standard deviations in parenthese) in the face match and the face mismatch conditions across *Reporting Style* in the face-unfocsed and the face-focused trials, respectively. Also shown are the paired-sample *t*-test results examining the simple effects of *Reporting Style* and *Face Match* in the two subsets of data.

Chapter 8 Conclusions and closing remarks

The experiments discussed in this thesis provide empirical support for the demonstration theory of direct speech (Clark & Gerrig, 1990) and embodied theories of language comprehension (e.g., Barsalou, 1999). In these experiments, we examined the mental representations during language comprehension of direct versus indirect speech. Specifically, we predominantly investigated whether individuals would be more likely to mentally simulate the reported speakers' voices during comprehension of direct speech as opposed to indirect speech as well as the nature of the simulated voice representations. Moreover, we explored whether representations of facial expressions are also mentally simulated during silent reading of direct versus indirect speech.

Experiments 1 and 2 (Chapters 2 and 3) obtained neuroimaging evidence of the hypothesised mental simulations of voices during language comprehension of direct speech rather than indirect speech. Experiment 1 demonstrated that silent reading of direct speech elicited higher brain activity within the voice-selective areas of the right auditory cortex as compared to silent reading of indirect speech. In the absence of external auditory stimulation, the observed brain activity within the voice-selective areas of the auditory cortex was interpreted as 'top-down' mental simulations of voices. Experiment 2 demonstrated that virtually the same brain areas of the auditory cortex became more active while listening to monotonously spoken direct as opposed to indirect speech. Such differential brain activations could not be accounted for by the acoustic characteristics or the vividness of the speech utterances without context other than the perceived congruency of the speech utterances in consideration of linguistic context. The findings suggest that listeners routinely expect more vivid vocal depictions for direct than for indirect speech. The findings are consistent with the notion that listeners spontaneously simulate supra-segmental acoustic representations when listening to the former in which such acoustic information is minimised. The observed brain activations in both experiments were located within virtually the same brain areas, including the posterior, middle and the anterior parts of the right superior temporal sulcus (STS). Taken together, we conclude from the two studies that language comprehension of direct speech is grounded in mental simulations of supra-segmental acoustic representations of voices whereas language comprehension of indirect speech is not, and that the reporting style of direct speech is intrinsically more vivid than that of the indirect speech. We also propose that the posterior,

middle and anterior parts of the right STS brain areas are involved in simulating such supra-segmental acoustic representations.

Experiments 3 to 10 (Chapters 4 to 6) sought to determine the behavioural consequences of the observed voice simulations during comprehension of direct versus indirect speech. These experiments examined contextual modulation of the acoustic dimensions (including pitch, loudness and rate) of the simulated voice representations during comprehension of direct and indirect speech. We found that *oral* readers spontaneously adjusted their own vocal acoustics (i.e., pitch, loudness or rate) in accordance with those of the reported speakers that were linguistically implied in the preceding contexts only when reading direct speech. The findings suggest that the simulated voice representations in comprehension of direct rather than indirect speech may be characterised in acoustic parameters of pitch, loudness and rate, and that voice simulations are not only reflected in brain-activation patterns but also in articulation. Moreover, we found that *silent* readers also adjusted their reading rates in accordance with the contextually implied speaking rates of the reported speakers when reading direct speech but not indirect speech. The results verify that speaking rate is indeed an important acoustic aspect of the simulated voice representations in comprehension of direct rather than indirect speech, and that voice simulations can also be reflected in eye-movement patterns. However, it remained inconclusive whether and how pitch and loudness are encoded in voice simulations during silent reading of direct rather than indirect speech. This could be because that the probe word was presented when reading the final word of the speech sentence (where voice simulations may have weakened over time), rendering the paradigm not sensitive enough to detect the simulated pitch or loudness representations. Future research should use a probe word that occurs early during reading of the speech utterance as opposed to at the final word of the sentence. Overall, we may conclude that the voice representations that are simulated in comprehension of direct rather than indirect speech can be characterised in modulations of speaking rate while further investigations are needed to illuminate how pitch and loudness are encoded in voice simulations.

To demonstrate a *multi-sensory* representation of direct speech, Experiment 11 (Chapter 7) ventured beyond voice simulations by examining whether individuals also mentally simulate, in addition to voices, the reported speakers' facial expressions during silent reading of direct rather than indirect speech. We assessed the effects of the implied facial

expressions of the reported speakers on the recognition latencies to graphic emoticons in which the illustrated facial expressions were made relevant by the previous text passage. We investigated three facial expressions but found that the 'surprised' expression was frequently confused with the 'happy' and 'angry' expressions. After excluding 'surprised'-related trials, we observed that when the linguistic focus was not on facial expressions, readers responded significantly faster to expression-congruent graphic emoticons than to expression-incongruent graphic emoticons after reading direct rather than indirect speech. This interaction was selectively masked when linguistic focus was placed on facial expressions, confirming that it reflected mental simulations of *facial expressions* (as opposed to *non-face* representations) during silent reading of direct rather than indirect speech. Future research is needed to replicate the current findings with more stimuli, focusing only the 'happy' and the 'angry' expressions.

In conclusion, this doctoral research showed that individuals are more likely to mentally simulate enriched supra-segmental acoustic representations of reported speakers' voices during language comprehension of direct speech as opposed to indirect speech. One aspect of such simulations may be characterised in modulations of speaking rate, but it remains unclear whether they encode acoustic representations of pitch or loudness. Moreover, it showed that representations of facial expressions may also be mentally simulated during silent reading of direct rather than indirect speech, as part of a multi-sensory representation of the former. The findings support the demonstration theory of direct speech (Clark & Gerrig, 1990) from the perspective of language comprehension, suggesting that the direct speech reporting style is intrinsically more vivid than the indirect speech reporting style. The results also extend embodied theories of language comprehension (e.g., Barsalou, 1999) in several respects. First of all, by demonstrating mental simulations of voices in language comprehension, this research extends the empirical testing of embodied theories to the auditory perceptual modality and to the sentence/discourse level which so far has not been thoroughly investigated in the relevant literature (previous studies have only focused on sound-related words, see Kellenbach et al., 2001; Kiefer et al., 2008). In this way, it contributes to the demonstration of the full spectrum of embodied phenomena in language comprehension. Second, while the contemporary empirical foundation of embodied language comprehension focuses on the grounding of the semantic meanings of language, the current research provides the first demonstration that language pragmatics, (e.g., the

vividness distinction between semantically-equivalent direct and indirect speech reporting styles) underlie language comprehension: In verbal communication, direct speech usually coincides with vivid paralinguistic demonstrations of the reported speech act whereas indirect speech is reported in a less vivid fashion. This research shows that this vividness distinction is also reflected in how language is processed, and that direct speech is mentally represented in vivid paralinguistic demonstrations and is more likely to be associated with mental simulations of voices than indirect speech. Third, this research combines event-related fMRI with eye-tracking to investigate neural correlates during on-line language comprehension; it revealed that the posterior, middle and anterior parts of the right STS are involved in mental simulations of supra-segmental acoustic representations. Such methods could inspire new applications in psycholinguistic research on language comprehension, helping us to understand the kinds of mental representations activated during reading.

At the end of the thesis, let us take a look into the future of embodied theories and their applications in language research. Embodied theories have marked a fundamental change in our current thinking about cognition by hypothesising that language, as well as other high-level cognitive functions, are 'embodied' in the sense that they are grounded in our sensory and motor experiences (implemented via some form of simulations/re-enactments) rather than in amodal symbols (e.g., Barsalou, 1999). However, contemporary empirical testing has been limited to *demonstrating* embodiment-type effects (e.g., interaction between linguistic and motoric tasks) in a rather superficial manner. To extend the scope of this empirical demonstration of embodied language phenomena, the current doctoral research combines both the traditional and the state-of-the-art neuroimaging methodologies and probes into both the behavioural consequences and the neural basis of the purported voice simulation processes. Nevertheless, this research cannot answer whether embodied simulations are functionally *necessary* or *causal* to language comprehension, or whether they play a relatively *peripheral*, or even *epiphenomenal*, role in language comprehension. Nor does it address how *abstract* language (e.g., the concept *democracy*) is mentally represented. These two issues (i.e., the necessity/causality of simulations in language comprehension and the representational nature for abstract language use) are often used to serve the argument against embodied theories, given that they have been the central issues in the development of embodied theories and the limitation of the contemporary empirical testing of these theories (Barsalou, 2010; Kiefer & Pulvermüller, in press). That said, the

current doctoral research provides a promising platform to address these issues. For example, to assess the necessities of voice simulation in comprehension of direct vs. indirect speech, we can use repetitive transcranial magnetic stimulation (rTMS) to temporarily disrupt the functions of the 'voice simulation'-related brain areas. The effects of such rTMS will implicate the causal role of voice simulations in language comprehension of direct vs. indirect speech. To explore the abstraction of voice simulations, one may be interested in investigating how deaf people, who have no experience of voice, mentally represent voices during comprehension of direct vs. indirect speech.

Based on the findings obtained within this doctoral thesis, future research has the potential to venture beyond demonstrations of embodied phenomena by seeking the evidence necessary to develop explicit mechanistic models of embodiment, using a variety of techniques such as eye-tracking, fMRI, diffusion tensor imaging (DTI), magnetoencephalography (MEG) and rTMS. The combination of these techniques can help us, for example, delineate the neural circuitry underlying simulations during written discourse comprehension, verify the necessity of voice simulations and explore the abstraction of voice simulations in the deaf population. The precise neural dynamics and verified properties of simulations will lay the biological and empirical foundation for sophisticated computational modelling of language comprehension and cognition, leading to development of fully intelligent systems and effective robots.

Appendices

Appendix 1(Stimuli for Experiment 1 & 2)

[D] = direct speech; [I] = indirect speech; [Q] = comprehension question

1. Julie and Mark had been classmates and have not seen each other for years. Today, they met in the local supermarket and Julie started a conversation about career paths.
[D] She said: 'My life has been amazing! After merely three years, I'm now a solicitor.'
[I] She said that her life had been amazing, and that after merely three years, she now was a solicitor.

2. One of Melanie's students, Jason, came into her office and said he could not reach her this morning. Melanie was confused because she had been in her office the whole time.
[D] She said: 'Well, in that case, there must be something wrong with my telephone.'
[I] She said that in that case, there must be something wrong with her telephone. [Q] Did Jason try to phone Melanie in the morning?

3. A Blackwell book store has recently opened in Edinburgh. Today, Alexia, a young mother living nearby, came in and asked for advice.
[D] She said: 'It's my son's birthday tomorrow and I would like to purchase some storybooks on adventure.'
[I] She said that it was her son's birthday soon and she would like to purchase some storybooks on adventure.

4. It was 5.30 pm and everybody was ready to leave the office. At one desk, Elaine was having a brief chat with Steven about her work.
[D] She said: 'Gosh, the amount of admin is killing me at the moment. I feel completely exhausted.'
[I] She said that the amount of admin was killing her at the moment, and that she felt completely exhausted.

5. It was February 14. Carolyn and Tony were on a date at the newly opened Chinese restaurant. The past couple of weeks, Carolyn was desperately trying to lose some weight.
[D] So she said to Tony: 'Oh, I hope you don't mind if I'm just having a starter.'
[I] So she said to Tony that hopefully he wouldn't mind if she were just having a starter.

6. Ramona and Keith were postgraduate students in Linguistics. Today, they met in the local café and Ramona started chatting about her favourite subjects.
[D] She said: 'Latin is boring! I'm much more interested in Eastern languages.'
[I] She said that she found Latin boring and that she was much more interested in Eastern languages.

7. Derek's birthday was just a couple of days away and his girlfriend Ruby had arranged something special. Tonight Ruby decided to unveil the plan that she kept hidden from him for so long.
[D] Over dinner, she said to Derek: 'Well, your birthday is coming up soon, so I booked us tickets to London to visit the opera.'
[I] Over dinner, she revealed to Derek that since his birthday was coming up soon, she booked them tickets to London to visit the opera.
[Q] Was Ruby's birthday coming up soon?

8. At the party on Friday night, some new faces were to be seen. Clare immediately caught Justin's eye, which went over to her and offered her a drink.
[D] Clare replied: 'Thanks very much, but I actually don't drink any alcohol.'
[I] Clare thanked him and mentioned that she actually doesn't drink any alcohol.
[Q] Did Justin catch Clare's eye?

9. Alison and Nick travelled to Beijing during the Olympics. Alison was amazed by the fine fabric on offer when they entered a big silk market.
[D] She said to Nick: 'I could buy the whole lot of it! The silk feels so incredibly smooth.'
[I] She said to Nick that she could buy the whole lot of it, and that the silk felt so incredibly smooth.

10. In the office, Ned told Daniela that his car had been damaged during an accident the other day. Daniela was sorry to hear that and tried to comfort him.
[D] She said: 'Well, at least you should get some money back from the insurance.'
[I] She said that at least he should get some money back from the insurance.
[Q] Had Ned's car been damaged?

11. At the department store, Colleen was busy working. She was serving a customer who wanted to know where he could buy some cosmetics for his wife.
[D] She replied: 'Ah, cosmetics are actually on the second floor. Some of them are at a discount.'
[I] She replied that cosmetics were actually on the second floor, adding that some of them were at a discount.

12. Jenny and Irvin were about to start a holiday trip to Barcelona. When they arrived at the airport to check in, Jenny noticed that a heavy thunderstorm was brewing outside.
[D] She complained: 'That's so unfair! I'm sure the plane will be delayed.'
[I] She complained that this was so unfair because she was sure that the plane would be delayed.

13. Britney is a student at the University of Glasgow. After a heavy snow in the afternoon, she was complaining to her boyfriend James about the weather.
[D] She said: 'I really hate the winter! It's always dark and the roads are too slippery.'
[I] She said that she really hated the winter because it's always dark and the roads are too slippery.

14. Jasmine had been ill for a week so her brother Colin took her to the doctor. After an hour, Jasmine came back to the waiting room looking rather disgruntled.
[D] She complained: 'The treatment was absolutely useless! Next time, I'd rather go to a specialist.'
[I] She complained that the treatment was absolutely useless and that, next time, she'd rather go to a specialist.

15. A famous symphony was going to be played for free at the university concert hall. Kate and Andrew were reading the advert, and Kate looked particularly interested.
[D] She said: 'It sounds really good and it's for free! I'm definitely going to that concert.'
[I] She said that it sounded really good and that it was for free, adding that she would definitely go to that concert.

16. On Valentine's Day, Simon at long last worked up the courage to propose to Emma. Emma wasn't sure, but at the same time, she didn't want to upset him.
[D] So she explained: 'You are a really nice bloke, Simon, but I'm simply not ready for such a commitment.'
[I] So she explained that although she thought he was a really nice bloke, she was simply not ready for such a commitment.

17. Two college students were talking about on-line shopping. Fraser thought it was pointless and boring, but Brenda couldn't disagree more.
[D] She said: 'What!? On-line shopping is absolutely amazing! I just snapped up some really cheap DVDs on Amazon.'
[I] She said she found on-line shopping absolutely amazing, and that she had just snapped up some really cheap DVDs on Amazon.
[Q] Did Fraser find on-line shopping boring?

18. It was quarter past seven and the Aberdeen Stock House was serving dinner. Debbie and Mike were very pleased with the meal and Debbie in particular was totally smitten.
[D] She said: 'We have to come here more often. The fish was really tasty and the soup was absolutely delicious.'
[I] She said that they should go there more often because the fish was really fresh and the soup was absolutely delicious.

19. It was 6-year-old Herbert's first day at school. His mother Laura, who had never let him play with other children before, was very anxious.
[D] She said to her husband: 'I'm really not sure whether Herbert is ready for this. I'd rather wait for another year.'
[I] She said to her husband that she was really not sure whether Herbert was ready for this, and that she'd rather wait for another year.

20. In order to finish the project in time, Audrey had been working in the office from 7am to 8pm without taking a break. Her colleague Sean just came back from dinner with a big smile on his face, which obviously upset her a bit.
[D] She complained: 'It's so unfair! My stomach has been rumbling all day. I could eat a whole elephant.'
[I] She complained that it was so unfair because her stomach had been rumbling all day, and that she could eat a whole elephant.

21. Cheryl and Barry were in the waiting lounge for their honeymoon flight to Paris. Shortly before boarding time, Cheryl noticed that her handbag was damaged.
[D] She shouted: 'Oh no! There's a hole in my bag! Don't tell me I've lost my passport!...'
[I] She shouted that there was a hole in her bag. She was worried that she had lost her passport.
[Q] Was Barry's bag broken?

22. It was eight o' clock in the evening when the Britain's Got Talent live recording session had finished. On their way out of the theatre, Sarah and Ross were having a giggle over some of the auditions.
[D] Sarah said: 'The jugglers at the end were funny, weren't they? Quite pathetic, actually, but very entertaining.'
[I] Sarah found that the jugglers at the end were very funny – quite pathetic, actually, but very entertaining.

23. Jessica and John were enjoying a hot and sunny day at the holiday resort in Spain. When they were heading towards their usual spot on the beach, Jessica was slightly overwhelmed by the heat.
[D] She said: 'Phew, this is almost a bit too hot for me! Perhaps we should first go back for a drink under the sunshade.'
[I] She said that it was almost a bit too hot for her, and suggested to first go back for a drink under the sunshade.

24. In the small town of Oban lies this charming little pub. One day, famous writer Aileen burst in with a big smile on her face.
[D] She said: 'Guess what – my new novel has been accepted for publication! I'll buy everyone a round.'
[I] She said that her new novel had been accepted for publication and that she wanted to buy everyone a round.

25. Theatrical agent Peter was visited by one of his clients. Shauna, a talented but struggling actress and singer, was desperate for a part in the new musical.
[D] She said: 'It's embarrassing to say, but without this job, I won't even be able to pay for electricity.'
[I] She was embarrassed to say that without this job, she wouldn't even be able to pay for electricity.
[Q] Was it Mary who visited Peter?

26. It was Sunday afternoon when housewife Heather was tidying up the bedroom. After just two minutes, she came out of the bedroom and confronted her husband Ben.
[D] She said: 'I can't believe how messy you are! The first thing I found on the floor was your dirty pyjamas!'
[I] She said that she could not believe how messy he was, and that the first thing she found on the floor was his dirty pyjamas.

27. Morton needed to speak to his project supervisor Marianne at her office. When he asked whether she could spare a minute, she seemed rather busy.
[D] She replied: 'I'm really sorry, but I have to finish marking some essays first.'
[I] She replied that she was really sorry and that she had to finish marking some essays first.

28. It was Christmas Eve and everyone was drinking and dancing at the local pub. Judith was already quite tipsy when she noticed that George was with a girl she had not seen before.
[D] She went over and complained to George: 'You've actually never told me that you had such a beautiful sister.'
[I] She went over and complained to George that he actually never told her that he had such a beautiful sister.
[Q] Was Judith George's sister?

29. Maureen had helped David a lot with his coursework, so David wanted to invite her for dinner. Asked what kind of food she would fancy, Maureen first thought for a second.
[D] Then she replied: 'I'm normally not that adventurous, but perhaps I should try some sushi.'
[I] Then she replied that although she would normally not be that adventurous, perhaps she should try some sushi.

30. Jane was a new post-graduate at the department. When she signed up at the Sports Club, she was having a conversation with Gregg about her favourite sports.
[D] She said: 'Oh, I haven't practised a lot recently, but I used to be quite good at tennis.'
[I] She said that although she had not practised a lot recently, she used to be quite good at tennis.

31. Leon and Betty were trying to get a mortgage for a new flat. While Betty was studying the relevant newspaper adverts, Leon was browsing some web-sites.
[D] After a while, he said to Betty: 'Hmm, these days, it looks as if nobody would give us an instant cashback.'
[I] After a while, he said to Betty that it looked as if nobody would give them an instant cashback.

32. Marcus and Helen were planning a trip to Europe for their wedding anniversary. Marcus was leaning more towards Prague while Helen preferred Rome.
[D] Marcus explained: 'Okay, Rome is nice, but the Czechs actually have much better beer.'
[I] Marcus explained that Rome would be nice, but that the Czechs would actually have much better beer.

33. Bill and Dianne went out bowling together. When Dianne wanted to pick a ball of her favourite colour, Bill looked a bit worried.
[D] He said: 'That one looks quite heavy. I'd rather pick the green one which I think is much smaller.'
[I] He said that the ball looked quite heavy and that he'd rather pick the green one which he thought was much smaller.
[Q] Was Dianne going to pick a heavy ball at first?

34. Rob and Becky were buying duvet covers at the store. When Becky chose a bright red one, Rob was completely against the colour.
[D] He said: 'No way! Just imagine that every morning you'd be waking up in a sea of blood!'
[I] He asked Becky just to imagine that every morning she would be waking up in a sea of blood.
[Q] Did Becky fancy a red duvet cover?

35. Harry and Leona were enjoying an afternoon walk in the park. When they were approaching the pond, Harry was impressed by this peaceful place.
[D] He said: 'This is so relaxing after a busy week! It is almost magical.'
[I] He found that it was so relaxing after a busy week, and that it was almost magical.

36. Schoolteacher Edward recently caused a stir by marrying Lynda, one of his ex-pupils. Annoyed by all the gossip, Edward told Lynda to simply not listen to what people say.
[D] He said: 'I think there's nothing wrong about a teacher marrying a former student.'
[I] He said that he thought there was nothing wrong about a teacher marrying a former student.
[Q] Had Lynda been Edward's teacher?

37. After making a small fortune in the lottery, Pamela bought quite an expensive car. Today in the car park, her colleague Ralph was extremely impressed by her new possession.
[D] He said: 'Wow! Is that your new car? It looks really stylish.'
[I] He asked her whether that was her new car, adding that it looked really stylish.
[Q] Was Pamela impressed by Ralph's new car?

38. Brian invited some of his mates home and started making a mess in the kitchen. His mum was obviously not happy about this, and Brian was quick to offer an excuse.
[D] He told her: 'Well, we just had a couple of drinks because the pubs are already closed.'
[I] He told her that they just had a couple of drinks because the pubs were already closed.

39. Daniel was injured in a car crash last week. At the hospital, he was quickly recovering. Today, he was surprised that some distant relatives came to visit him.
[D] Not without sarcasm, Daniel said: 'Oh, looks as if one has to get hurt to get some flowers.'
[I] Not without sarcasm, Daniel said that it looked as if one had to get hurt to get some flowers.

40. Jacob had promised Cindy to buy a bonsai tree to decorate their living room. When Cindy wanted to place the plant by the television, he tried to persuade her of an alternative option.
[D] He said: 'Look, if you do that, there will be no sunlight for the bonsai, so it's better to place it by the window.'
[I] He said that if she did that, there would be no sunlight for the bonsai, so it was better to place it by the window.

41. PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. When Ella asked for an extension, Jim looked concerned.
[D] He said: 'Hmm, we really need those data in by next month for that conference.'
[I] He said that they really needed those data in by next month for that conference.
[Q] Was Ella Jim's PhD student?

42. Albert was taking Joanne to the new Schwarzenegger movie. Albert was very excited, but since Joanne seemed not very keen, he tried to persuade her to go with him.
[D] He said: 'Come on! I'm sure you will love it! The movie has some really cool special effects.'
[I] He said he was sure she would love it, and that the movie had some really cool special effects.

43. Neil and Stephanie were visiting the local distillery. Neil was amazed by the range of different Whiskeys produced, and he tried a very expensive Single Malt.
[D] He said: 'Gee, that's a strong one! I'm glad I won't have to drive us home.'
[I] He said it was really strong and that he was glad he didn't have to drive them home.
[Q] Did Neil try a Single Malt?

44. Clive and his wife Molly visited the new seafood restaurant for the first time. Clive was extremely anxious about food poisoning and did not want to have anything raw.
[D] So he said to Molly: 'I don't care what other people might think, but I will order properly cooked oysters.'
[I] So he said to Molly that, regardless of what other people might think, he would order properly cooked oysters.

45. Luke and his friends were watching a movie at the cinema. Luke wasn't particularly keen on romantic comedies, and he was complaining a lot after the film.
[D] He said: 'God, that movie was terrible! I've never been so bored in my life.'
[I] He said that the movie was terrible and that he had never been so bored in his life.

46. On the train to London, Thomas was very nervous about his upcoming job interview. When he reached for a cigarette to calm himself down, the middle-aged man next to him complained instantly.
[D] The man called out: 'Oi! Smoking on public transport is illegal!'
[I] He called out that smoking on public transport was illegal.

47. During a coffee break, university teachers Robert and Isabel met in the common room. Robert was slightly irritated with the mess in the kitchen.
[D] He said to Isabel: 'Look at this! Certain people must always leave their dirty mugs on the table.'
[I] He said to Isabel that, obviously, certain people must always leave their dirty mugs on the table.
[Q] Was Isabel complaining about the dirty mugs?

48. School teachers Duncan and Liana were arguing about the curriculum. Apparently, Duncan was not a big fan of Liana's modern style of teaching.
[D] He said: 'If my own children were at this school, I'd make sure they would never have to attend your lessons.'
[I] He said that if his own children were at this school, he'd make sure they would never have to attend her lessons.

49. It was getting quite late and most of Andrea's friends were about to leave her party. Scott in particular seemed to have had a great time when he was leaving.
[D] He said: 'Sorry about the crack in that glass. The party was absolutely fantastic.'
[I] He apologised for the crack in that glass, adding that the party was absolutely fantastic.
[Q] Did Scott throw a party?

50. Doug always enjoyed a bit of gardening in his spare time. He recently bought a new lawnmower, but he was quite frustrated because it simply refused to work.
[D] He complained: 'Sod it! Things used to be much better in the past. All one can buy these days is rubbish!'
[I] He complained that things used to be much better in the past, and that all one can buy these days was rubbish.

51. After class, Roy had a serious conversation with one of his pupils. The boy had been accused of bullying and Roy was giving him a caution.
[D] Roy said: 'If the accusations are true, you will be in serious trouble. One cannot tolerate such kind of behaviour.'
[I] Roy said that if the accusations were true, the boy would be in serious trouble, and that one cannot tolerate such kind of behaviour.

52. In the board room, assistant manager Craig was giving a presentation on how to improve the company's sales figures. When he noticed that his boss was anything but impressed, Craig used his lack of experience as an excuse.
[D] He said: 'Sorry, this is actually my first ever presentation and I'm probably not very persuasive.'
[I] He said that this was actually his first ever presentation, and that he was probably not very persuasive.

53. Alice was working in a small antiques shop down the local high street. Today, a weird-looking man with greasy hair and thick glasses came into the shop.
[D] He looked around and said: 'You may be surprised to learn that I'm a world-renowned collector of rare pheasant paintings.'
[I] He looked around and said that she might be surprised to learn that he was a world-renowned collector of rare pheasant paintings.

54. Ray and Sandra were on a trip to the countryside for the weekend. Ray quite liked the outdoors and particularly that rural smell.
[D] He said: 'Aah, this is so much better than crowded cities and pollution.'
[I] He said that it was so much better than crowded cities and pollution.

55. Earlier in the afternoon, Max and Janis were having a cigarette in front of the main entrance. Not knowing what to say, Max started his typical conversation about the weather.
[D] He said: 'Gosh, it's quite windy today – but at least it's not raining.'
[I] He said that although it was quite windy today, at least it was not raining.

56. A journalist was interviewing Eric, the older brother of a famous pop diva. Eric did not really enjoy the attention he was given and wasn't sure what to say.
[D] He proclaimed: 'I actually don't see my sister very often these days. She is completely devoted to her career.'
[I] He proclaimed that he actually did not see his sister very often in those days, and that she was completely devoted to her career.

57. Thomas went to the local McDonalds and incidentally met Sheena, who he knew from former Weight Watchers meetings. Naturally, they were both quite embarrassed, but Thomas finally broke the silence.
[D] He said: 'Weight Watchers is rubbish, isn't it. I've tried so many diets but I'm still overweight.'
[I] He said that Weight Watchers was rubbish, and that he had tried so many diets but was still overweight.

58. A car-boot sale was taking place in the neighbourhood and Kenny was trying to get rid of some of the old records he owned since the 1970s. A trendy young woman appeared to be very interested, so Kenny started advertising.
[D] He said: 'I can give you three albums for a fiver. You won't find any of that stuff on the internet.'
[I] He said that he could give her three albums for a fiver, adding that she wouldn't find any of that stuff on the internet.

59. A trendy night club had been recently opened near Marvin's flat and it proved instantly popular. However, Marvin often felt disturbed by the noise, and this morning he was complaining to the neighbours about it.
[D] He said: 'This is unbelievable! Every time I have to work early hours, they are having a bash!'
[I] He said that every time he has to work early hours, they are having a bash.

60. Carl, a taxi driver, was about to drive home when suddenly he saw a man by the road frantically waving his arms. Carl stopped and the man offered him double the fare if they make it to the airport within ten minutes.
[D] Carl replied: 'I'll try my best, mate, but cannot work any miracles.'
[I] Carl replied that he would try his best, but that he could not work any miracles.

61. It was 11am in the morning when the fire alarm went off. Hearing people running down the corridors, Mary grabbed her jacket and burst into Peter's office next door.
[D] She shouted: 'Peter, quick, we have to leave immediately because the building is on fire!'
[I] She urged Peter to leave immediately because the building was on fire.

62. It was early Tuesday morning when Claire was looking for the train tickets she left in one of her handbags the other day. After searching for a couple of minutes she became worried and confused.
[D] She said to partner Gareth: 'Erm, that's weird... I am pretty sure they were in that black leather bag, but now they're gone.'
[I] She said to her partner Gareth that it was weird because she was pretty sure that they were in that black leather bag, but now they were gone.

63. Twenty-five year old Connie was going to sing in the quarter-final of a local talent competition. She was extremely nervous before her performance, even though her mother had tried to calm her down.
[D] She said: 'No! I can't do it! This is the end of the journey. I can't face the audience this time!'
[I] She said that she could not do it and that was the end of the journey because she couldn't face the audience this time.

64. At the beach, William and Phillip were enjoying the sunshine and a fresh sea breeze. Suddenly, William noticed a piece of wood floating in the sea, with something lying on it that looked like a lifeless human body.
[D] Without hesitation, William exclaimed: 'Phil, look! There's someone in trouble! We have to get some help quickly!'
[I] Without hesitation, William exclaimed that there was someone in trouble and that they had to get some help quickly. [Q] Was it William who found the lifeless human body?

65. At Mr. Harris's house, Police constables Paul and Ken had some sad news to impart. Mrs. Harris's husband, a well-known businessman, was killed in a car crash earlier this morning.
[D] When Paul spoke with her, he said: 'Mrs. Harris, I am afraid that your husband was involved in a road accident. He was announced dead at the scene.'
[I] When Paul spoke with her, he said that her husband was involved in a road accident and that he was announced dead at the scene.

66. On the train to London, young mother Olivia was taking her 5 year old daughter for a holiday. The little girl could not understand why the youth who lit a cigarette at the station was scolded by other passengers.
[D] She asked: 'If people don't like smoking, why don't they just hold their breath?'
[I] She suggested that if people don't like smoking, they could just hold their breath.
[Q] Did the little girl think the youth could smoke at the station?

67. It was quarter past seven at the French restaurant, and head chef Pascal was extremely distressed. During a routine check, the restaurant manager went berserk about the messy state of the kitchen.
[D] He said: 'Is this how you want to make yourself a reputation?! It's a miracle that the health inspectors haven't closed us down yet!'
[I] He asked whether this would be the way he wanted to make himself a reputation, adding that it's a miracle that the health inspectors haven't closed them down yet.

68. Teenagers Helen and Laura were shopping around at the summer sale. Helen was very fond of a black skirt and shrieked when she found that it was heavily discounted.
[D] She called out: 'Laura, look! This is exactly what I always wanted, and it's merely fifty quid!'
[I] She called out to Laura that this is exactly what she always wanted, and that it was merely fifty quid.

69. On Sunday afternoon, the Riley's were going to the train station to pick up granddad who lived in the neighbouring town. Twelve year-old Karen did not understand why they went so early if the train was arriving at 6 pm.
[D] Mr Riley replied: 'Well, just in case.. Remember last time when granddad forgot to set his watch properly?'
[I] Mr Riley replied that this was just in case, reminding her of the last time when granddad forgot to set his watch properly.

70. A medical emergency tutorial was taking place in the operating theatre. The tutor was demonstrating the steps to treat a heavy injury to the students.
[D] He explained: 'When the patient comes in, first set up the life-support machine and then stop the bleeding.'
[I] He explained that when the patient comes in, they first should set up the life-support machine and then stop the bleeding.
[Q] Is it correct to stop the bleeding first?

71. At Omaha Beach, allied soldiers braved artillery fire, mortar attacks and all-round danger. In response to the enemy's gunfire, the captain of the 1st Division rallied the soldiers.
[D] He shouted: 'Be brave, lads! Toss your grenades and open fire with your rifles! We must silence that machine gun.'
[I] He told the soldiers to toss their grenades and open fire with their rifles, for they must silence that machine gun.

72. The West End festival was coming soon and the organisers were discussing details of the live television broadcast. The current issue was where to set the television cameras and John was offering his plans.
[D] He said: 'We might need television cameras along the High Street, and possibly even a helicopter to cover the whole area.'
[I] He suggested that they might need television cameras along the High Street and possibly even a helicopter to cover the whole area.

73. Today in Northern Afghanistan, the United Nations Peacekeeping Force was ambushed by Taliban forces. Although they returned fire immediately, the UN troops suffered heavy losses.
[D] The officer hastily called for reinforcements: 'We are facing rocket and mortar fire in the region! We urgently need armoured support!'
[I] The officer hastily called for reinforcements, confirming that they were facing rocket and mortar fire in the region and that they urgently needed armoured support.

74. It was midnight already and moonlight dripped through the leaves on the pavements. Robert and Chloe just came back from a party, and Chloe was still full of energy.
[D] She patted Robert's cheek and said: 'Stop yawning, my darling! The day is far from over yet.'
[I] She patted Robert's cheek and said that he should stop yawning because the day was far from over yet.

75. Joseph, an extremely wealthy entrepreneur, was known to be addicted to horse racing. Tipped-off by an insider, he placed a huge bet on one horse today, which failed to win the race, however.
[D] Joseph was now shouting furiously on the phone: 'Where did your bloody information come from!? I've lost nearly one million pounds!'
[I] Joseph was now shouting furiously on the phone, asking where the information had come from, and claiming that he had lost nearly one million pounds.
[Q] Was Joseph confident about the bet?

76. Due to the economic recession, Ron's company has lost a great quantity of international trades. His wife noticed that every day he just slumped on the couch, with a tired and dull look on his face.
[D] She tried to comfort him by saying: 'Why don't you stop your business for a while? We could spend more time together, and there is certainly a positive side to that!'
[I] She tried to comfort him by suggesting that if he stops his business for a while, they could have more time together, and there was certainly a positive side to that.

77. At the airport, Carol and Fraser were boarding the flight to Barcelona for a conference. There were only 5 minutes left before the plane was taking off, and Carol started panicking because she couldn't find her boarding card.
[D] She turned to Fraser and said: 'Oh no! I can't find my stupid boarding pass! I must have left it at the duty-free shop!'
[I] She turned to Fraser and said that she couldn't find her boarding pass and that she must have left it at the duty-free shop.

78. At Glasgow Royal Infirmary, family members were sitting around a bed, feeling sad. A very old man was dying, and too weak to sit up. He wanted to say something, so his daughter placed a cushion under his head.
[D] He looked around slowly, thanked them for their coming, and said: 'I am so happy to have the whole family around me in my final hour.'
[I] He looked around slowly, thanked them for their coming, and added that he was so happy to have the whole family around him in his final hour.

79. Gerry and Herbert were attending a first aid tutorial where they were practising resuscitation procedures using a plastic dummy. Gerry was busy remembering all the steps, but Herbert wouldn't bother and instead explained his own strategy.
[D] He said: 'Look, it's not necessary to go through all the steps. The important thing is to ensure that air passages are not obstructed.'
[I] He said that it's not necessary to go through all the steps, and that the important thing was to ensure that air passages were not obstructed.

80. At House of Fraser, the manager was demonstrating her employees how to describe, in one sentence, the special features of a product. She did it by performing a mock sales pitch in front of them.
[D] Holding up a jacket, she said: 'This leather jacket is made of the finest materials and represents the latest Italian couture.'
[I] Holding up a jacket, she said that it was made of the finest materials and represented the latest Italian couture.
[Q] Was the leather jacket from France?

81. A heavily armed robbery took place at the jewelleries' early this morning. The criminals escaped in a blue van. Within five minutes, three police cars, an armed response unit, and an ambulance showed up at the scene.
[D] The officer-in-charge ordered: 'Every main road must be blocked immediately! And, for Christ's sake, call the helicopter in!'
[I] The officer-in-charge ordered that every main road must be blocked immediately, and that the helicopter should be called in.

82. David was producing a TV documentary about the Vietnam War, and today, he was interviewing a 75 year-old Vietnam veteran. When David asked about the situation at the battle of Hue, the old soldier paused for a while.
[D] Then, slowly, he said: 'This battle was one of the longest and most brutal of the entire war. It's haunting me in my dreams for the rest of my life.'
[I] Then, slowly, he said that this battle was one of the longest and most brutal of the entire war, and that it would be haunting him in his dreams for the rest of his life.

83. It was a Sunday in June, and a young couple were waiting for their relatives to arrive. Zoe was busy making cucumber sandwiches and scones, when her husband noticed that her hands were actually shaking a bit.
[D] She explained: 'Don't worry, I'm OK! I'm just a little nervous because I really want everything to be perfect.'
[I] She explained that she was OK, and that she was just a little nervous because she really wanted everything to be perfect.

84. University lecturer Bridget was responsible for admission to the English Literature classes. This morning, she had to deal with yet another request from a student who wanted to change her course.
[D] The student said: 'I'm not sure whether Contemporary Women's Poetry is right for me. Is it possible to switch to Sexual Identity in the Works of Oscar Wilde instead?'
[I] The student said that she was not sure whether Contemporary Women's Poetry was right for her and asked if it was possible to switch to Sexual Identity in the Works of Oscar Wilde instead.

85. In the computer shop, a customer wanted to get his laptop fixed because the Windows upgrade hadn't installed correctly. The shop assistant had a quick look and noticed that the customer was using the Home Edition.
[D] He said: 'Well, I presume you've tried the wrong update. You can't really apply updates for XP Professional to the Home Edition.'
[I] He presumed that the customer had tried the wrong update, adding that one cannot really apply updates for XP Professional to the Home Edition.

86. At Saint Jacob's hospital in Falmouth, Henry was waiting to have an X-ray taken. He fell off the bus this morning and feared he broke his leg, but worst of all, he was also very anxious about X-rays.
[D] The doctor said: 'No need to worry! We will give you a lead board to protect you from radiation.'
[I] The doctor told him not to worry because they would give him a lead board to protect him from radiation.

87. The heat of the morning was intense but Charles and Lucy had managed to climb up a hill. From the top, Charles could see right down to the beach on the other side.
[D] Excited, he said to Lucy: 'Brilliant! Let's run down to the beach and take a plunge into the sea!'
[I] Excited, he suggested to Lucy that they should run down to the beach and take a plunge into the sea. [Q] Did Lucy suggest that they should take a plunge into the sea?

88. Mrs. Jones felt awkward to ask her son for help because she ran out of cash for the babysitter. She pretended to enter her son's room 'by accident', trying to borrow £20.
[D] The boy replied: 'You see, I'd love to help you out, but your credit history isn't what I would call spotless.'
[I] The boy replied that he would love to help her out, but that her credit history wasn't what he would call spotless. [Q] Did Mrs. Jones successfully borrow the money?

89. Eleanor's son had always been a bit of a couch potato. Over dinner, Eleanor was passing a cutting from a London newspaper to her husband.
[D] She said: 'There, just read this article about child education! It exactly confirms my concerns about watching too much TV.'
[I] She asked him to read that article about child education, for it exactly confirmed her concerns about watching too much TV.

90. The Smiths had recently suffered from some financial problems, so Mr. Smith was looking for a better job. Today, he was going to have an interview with the manager of a big insurance company.
[D] Before he left, his wife said: 'Best of luck, darling. Always remember that you are the best man for the job!'
[I] Before he left, his wife wished him best of luck, and advised him to remember that he is the best man for the job.

Appendix 2 Materials for oral reading (Experiment 3)

[D] = direct speech; [I] = indirect speech; [VF] = voice focused; [VU] = voice unfocused

Item	Female reported speaker	Male reported speaker
1	Julie and Mark had been classmates and have not seen each other for years.Today, they met in the local pub and Julie started a conversation about career paths. **[VF - D]** She was very pleased with her job and cheerfully talked about it, almost like a newly wedded woman, saying: 'My life has been amazing! After merely three years, I'm now a solicitor.' **[VF - I]** She was very pleased with her job and cheerfully talked about it, almost like a newly wedded woman, saying that her life had been amazing, and that after merely three years, she now was a solicitor. **[VU - D]** She was very pleased with her job and was gesturing a lot, almost like a 12-year-old teenager, saying: 'My life has been amazing! After merely three years, I'm now a solicitor.' **[VU - I]** She was very pleased with her job and was was gesturing a lot, almost like a 12-year-old teenager, saying that her life had been amazing, and that after merely three years, she now was a solicitor.	Mark and Julie had been classmates and have not seen each other for years.Today, they met in the local pub and Mark started a conversation about career paths. **[VF - D]** He was very pleased with his job and cheerfully talked about it, almost like a newly-wed, saying: 'My life has been amazing! After merely three years, I'm now a solicitor.' **[VF - I]** He was very pleased with his job and cheerfully talked about it, almost like a newly-wed, saying that his life had been amazing, and that after merely three years, he now was a solicitor. **[VU - D]** He was very pleased with his job and was gesturing a lot, almost like a 12-year-old teenager, saying: 'My life has been amazing! After merely three years, I'm now a solicitor.' **[VU – I]** He was very pleased with his job and was was gesturing a lot, almost like a 12-year-old teenager, saying that his life had been amazing, and that after merely three years, he now was a solicitor.
2	One of Melanie's students, Jason, came into her office and said he could not reach her this morning. Melanie could not believe it because she had been in her office the whole time. **[VF – D]** She replied with a sarcastic tone in her voice: 'Really?! Well, in that case, there must be something wrong with my telephone.' **[VF – I]** She replied with a sarcastic tone in her voice, saying that in that case, there must be something wrong with her telephone. **[VU – D]** She picked up the phone and listened carefully, concluding: 'Well, in that case, there must be something wrong with my telephone.' **[VU – I]** She picked up the phone and listened carefully, and concluded that in that case, there must be something wrong with her telephone.	One of Jason's students, Melanie, came into his office and said she could not reach him this morning. Jason could not believe it because he had been in his office the whole time. **[VF – D]** He replied with a sarcastic tone in his voice: 'Really?! Well, in that case, there must be something wrong with my telephone.' **[VF – I]** He replied with a sarcastic tone in his voice, saying that in that case, there must be something wrong with his telephone. **[VU – D]** He picked up the phone and listened carefully, concluding: 'Well, in that case, there must be something wrong with my telephone.' **[VU – I]** He picked up the phone and listened carefully, and concluded that in that case, there must be something wrong with his telephone.
3	A Blackwell book store has recently opened in Edinburgh.Today, Alexia, a young mother living nearby, came in and looked for a present for her little boy. **[VF – D]** Her voice was chirpy and melodic, saying: 'I'm very excited! It's my son's birthday tomorrow and I want to surprise him with his favourite storybooks on adventure.' **[VF – I]** Her voice was chirpy and melodic, saying that she was very excited, for it was her son's birthday tomorrow and she wanted to surprise him with his favourite storybooks on adventure. **[VU – D]** She said, with a sparkle in her eyes: 'I'm very excited! It's my son's birthday tomorrow and I want to surprise him with his favourite storybooks on adventure.' **[VU – I]** She said, with a sparkle in her eyes, that she was very excited, for it was her son's birthday tomorrow and she wanted to surprise him with his favourite storybooks on adventure.	A Blackwell book store has recently opened in Edinburgh.Today, Alex, a young father living nearby, came in and looked for a present for his little boy. **[VF – D]** His voice was chirpy and melodic, saying: 'I'm very excited! It's my son's birthday tomorrow and I want to surprise him with his favourite storybooks on adventure.' **[VF – I]** His voice was chirpy and melodic, saying that he was very excited, for it was his son's birthday tomorrow and he wanted to surprise him with his favourite storybooks on adventure. **[VU – D]** He said, with a sparkle in his eyes: 'I'm very excited! It's my son's birthday tomorrow and I want to surprise him with his favourite storybooks on adventure.' **[VU – I]** He said, with a sparkle in his eyes, that he was very excited, for it was his son's birthday tomorrow and he wanted to surprise him with his favourite storybooks on adventure.

4	It was 5.30 pm and everybody was ready to leave the office.At one desk, Elaine was having a brief chat with Steven about her work. **[VF – D]** She complained, with a hint of reluctance in her voice: 'Gosh, the amount of admin is killing me at the moment. I feel completely exhausted!' **[VF – I]** She complained, with a hint of reluctance in her voice, that the amount of admin was killing her at the moment, and that she felt completely exhausted. **[VU – D]** She complained, with her face looking pale and tired: 'Gosh, the amount of admin is killing me at the moment. I feel completely exhausted.' **[VU – I]** She complained, with her face looking pale and tired, that the amount of admin was killing her at the moment, and that she felt completely exhausted.	It was 5.30 pm and everybody was ready to leave the office.At one desk, Steven was having a brief chat with Elaine about his work. **[VF – D]** He complained, with a hint of reluctance in his voice: 'Gosh, the amount of admin is killing me at the moment. I feel completely exhausted!' **[VF – I]** He complained, with a hint of reluctance in his voice, that the amount of admin was killing him at the moment, and that he felt completely exhausted. **[VU – D]** He complained, with his face looking pale and tired: 'Gosh, the amount of admin is killing me at the moment. I feel completely exhausted.' **[VU – I]** He complained, with his face looking pale and tired, that the amount of admin was killing him at the moment, and that he felt completely exhausted.
5	It was February 14. Carolyn and Tony were on a date at the newly opened Chinese restaurant. The past couple of weeks, Carolyn was desperately trying to lose some weight. **[VF – D]** She didn't want to disappoint Tony, so she said to him, softly and gently: 'Oh darling, I hope you don't mind if I'm just having a starter.' **[VF – I]** She didn't want to disappoint Tony, so she said to him, softly and gently, that hopefully he wouldn't mind if she was just having a starter. **[VU – D]** She didn't want to disappoint Tony, so she said to him, with an apologetic smile: 'Oh darling, I hope you don't mind if I'm just having a starter.' **[VU – I]** She didn't want to disappoint Tony, so she said to him, with an apologetic smile, that hopefully he wouldn't mind if she was just having a starter.	It was February 14. Tony and Caroline were on a date at the newly opened Chinese restaurant. The past couple of weeks, Tony was desperately trying to lose some weight. **[VF – D]** He didn't want to disappoint Caroline, so he said to her, softly and gently: 'Oh darling, I hope you don't mind if I'm just having a starter.' **[VF – I]** He didn't want to disappoint Caroline, so he said to her, softly and gently, that hopefully she wouldn't mind if he was just having a starter. **[VU – D]** He didn't want to disappoint Caroline, so he said to her, with an apologetic smile: 'Oh darling, I hope you don't mind if I'm just having a starter.' **[VU – I]** He didn't want to disappoint Caroline, so he said to her, with an apologetic smile, that hopefully she wouldn't mind if he was just having a starter.
6	Ramona and Keith were postgraduate students in Linguistics. Today, they met in the local café and Ramona started chatting about their favourite subjects. **[VF – D]** When the topic switched to Latin, she sounded really upset with a moaning voice: 'Latin is so boring! I'm much more interested in Eastern languages.' **[VF – I]** When the topic switched to Latin, she sounded really upset with a moaning voice, saying that she found Latin so boring and that she was much more interested in Eastern languages. **[VU – D]** When the topic switched to Latin, she had a really upset look on her face: 'Latin is so boring! I'm much more interested in Eastern languages.' **[VU – I]** When the topic switched to Latin, she had a really upset look on her face, saying that she found Latin so boring and that she was much more interested in Eastern languages.	Keith and Ramona were postgraduate students in Linguistics. Today, they met in the local café and Keith started chatting about their favourite subjects. **[VF – D]** When the topic switched to Latin, he sounded really upset with a moaning voice: 'Latin is so boring! I'm much more interested in Eastern languages.' **[VF – I]** When the topic switched to Latin, he sounded really upset with a moaning voice, saying that he found Latin so boring and that he was much more interested in Eastern languages. **[VU – D]** When the topic switched to Latin, he had a really upset look on his face: 'Latin is so boring! I'm much more interested in Eastern languages.' **[VU – I]** When the topic switched to Latin, he had a really upset look on his face, saying that he found Latin so boring and that he was much more interested in Eastern languages.

7	Derek's birthday was just a couple of days away and his girlfriend Ruby had arranged something special. Tonight Ruby decided to unveil the plan that she kept hidden from him for so long. **[VF- D]** Over dinner, she drew Derek's attention with a chirpy tone in her voice: 'Look, darling, your birthday is coming up soon, so... I booked us tickets to London to visit the opera!' **[VF – I]** Over dinner, she drew Derek's attention with a chirpy tone in her voice, saying that since his birthday was coming up soon, she booked them tickets to London to visit the opera. **[VU – D]** Over dinner, she drew Derek's attention with luring gesture: 'Look, darling, your birthday is coming up soon, so... I booked us tickets to London to visit the opera.' **[VU – I]** Over dinner, she drew Derek's attention with a luring gesture, saying that since his birthday was coming up soon, she booked them tickets to London to visit the opera.	Ruby's birthday was just a couple of days away and her boyfriend Derek had arranged something special. Tonight Derek decided to unveil the plan that he kept hidden from her for so long. **[VF – D]** Over dinner, he drew Ruby's attention with a chirpy tone in his voice: 'Look, darling, your birthday is coming up soon, so... I booked us tickets to London to visit the opera!' **[VF – I]** Over dinner, he drew Ruby's attention with a chirpy tone in his voice, saying that since her birthday was coming up soon, he booked them tickets to London to visit the opera. **[VU – D]** Over dinner, he beckoned Ruby over: 'Look, darling, your birthday is coming up soon, so... I booked us tickets to London to visit the opera.' **[VU – I]** Over dinner, he beckoned Ruby over, saying that since his birthday was coming up soon, she booked them tickets to London to visit the opera.
8	At the party on Friday night, some new faces were to be seen. Clare immediately caught Justin's eye, who went over to her and offered her a drink. **[VF – D]** 'Oh?!' she sounded a little flatterred: 'Thanks very much, but I actually don't drink any alcohol.' **[VF – I]** Clare sounded a little flattered. She thanked him and mentioned that she actually doesn't drink any alcohol. **[VU – D]** 'Oh?!' Clare smiled like a blooming sunflower, 'Thanks very much, but I actually don't drink any alcohol.' **[VU – I]** Clare smiled like a blooming sunflower. She thanked him and mentioned that she actually doesn't drink any alcohol.	At the party on Friday night, some new faces were to be seen. Justin immediately caught Claire's eye, who went over to him and offered him a drink. **[VF – D]** 'Oh?!' he sounded a little flattered: 'Thanks very much, but I actually don't drink any alcohol.' **[VF – I]** Justin sounded a little flattered. He thanked her and mentioned that he actually doesn't drink any alcohol. **[VU – D]** 'Oh?!' Justin smiled politely, 'Thanks very much, but I actually don't drink any alcohol.' **[VU – I]** Justin smiled politely. He thanked her and mentioned that he actually doesn't drink any alcohol.
9	Alison and Nick travelled to Beijing during the Olympics. Alison was amazed by the fine fabric on offer when they entered a big silk market. **[VF – D]** Her voice was vibrating frantically: 'Oh my goodness! I could buy the whole lot of it! The silk feels so incredibly smooth.' **[VF – I]** Her voice was vibrating frantically, saying that she could buy the whole lot of it, and that the silk felt so incredibly smooth. **[VU – D]** Her eyes were wide open, like two plump strawberries: 'Oh my goodness! I could buy the whole lot of it! The silk feels so incredibly smooth.' **[VU – I]** Her eyes were wide open, like two plump strawberries when she said that she could buy the whole lot of it, and that the silk felt so incredibly smooth.	Nick and Alison travelled to Beijing during the Olympics. Nick was amazed by the fine fabric on offer when they entered a big silk market. **[VF – D]** His voice was vibrating frantically: 'Oh my goodness! I could buy the whole lot of it! The silk feels so incredibly smooth.' **[VF – I]** His voice was vibrating frantically, saying that he could buy the whole lot of it, and that the silk felt so incredibly smooth. **[VU – D]** His eyes were wide open, like two plump strawberries: 'Oh my goodness! I could buy the whole lot of it! The silk feels so incredibly smooth.' **[VU – I]** His eyes were wide open, like two plump strawberries when he said that he could buy the whole lot of it, and that the silk felt so incredibly smooth.

10	In the office, Ned told Daniela that his car had been damaged during an accident the other day. Daniela was sorry to hear that and tried to comfort him. **[VF –D]** She said in a calming and reassuring tone: 'Well, at least you should get some money back from the insurance.' **[VF – I]** She said in a calming and reassuring tone: that at least he should get some money back from the insurance. **[VU – D]** She said, with her hand gently patting his back: 'Well, at least you should get some money back from the insurance.' **[VU – I]** She said, with her hand gently patting his back, that at least he should get some money back from the insurance.	In the office, Daniela told Ned that her car had been damaged during an accident the other day. Ned was sorry to hear that and tried to comfort her. **[VF – D]** He said in a calming and reassuring tone: 'Well, at least you should get some money back from the insurance.' **[VF – I]** He said in a calming and reassuring tone: that at least she should get some money back from the insurance. **[VU – D]** He said, with his hand gently patting her back: 'Well, at least you should get some money back from the insurance.' **[VU – I]** He said, with his hand gently patting her back, that at least she should get some money back from the insurance.
11	At the department store, Colleen was busy working. She was serving a customer who wanted to know where he could buy some cosmetics for his wife. **[VF – D]** She replied in a polite and clear voice: 'Ah, cosmetics are actually on the second floor. Some of them are at a discount.' **[VF – I]** She replied in a polite and clear voice that cosmetics were actually on the second floor, adding that some of them were at a discount. **[VU – D]** She replied with a professional smile: 'Ah, cosmetics are actually on the second floor. Some of them are at a discount.' **[VU – I]** She replied with a professional smile that cosmetics were actually on the second floor, adding that some of them were at a discount.	At the department store, Colin was busy working. He was serving a customer who wanted to know where she could buy some cosmetics for her friend. **[VF – D]** He replied in a polite and clear voice: 'Ah, cosmetics are actually on the second floor. Some of them are at a discount.' **[VF – I]** He replied in a polite and clear voice that cosmetics were actually on the second floor, adding that some of them were at a discount. **[VU – D]** He replied with a professional smile: 'Ah, cosmetics are actually on the second floor. Some of them are at a discount.' **[VU – I]** He replied with a professional smile that cosmetics were actually on the second floor, adding that some of them were at a discount.
12	Jenny and Irvin were about to start a holiday trip to Barcelona. When they arrived at the airport to check in, Jenny noticed that a heavy thunderstorm was brewing outside. **[VF – D]** She became almost hysterical, squeaking: 'That's so unfair! I'm sure the plane will be delayed.' **[VF – I]** She became almost hysterical, squeaking that this was so unfair because she was sure that the plane would be delayed. **[VU – D]** She complained with a very disappointed look on her face: 'That's so unfair! I'm sure the plane will be delayed.' **[VU – I]** She complained, with a very disappointed look on her face, that this was so unfair because she was sure that the plane would be delayed.	Irvin and Jenny were about to start a holiday trip to Barcelona. When they arrived at the airport to check in, Irvin noticed that a heavy thunderstorm was brewing outside. **[VF – D]** He became almost hysterical, growling: 'That's so unfair! I'm sure the plane will be delayed.' **[VF – I]** He became almost hysterical, growling that this was so unfair because he was sure that the plane would be delayed. **[VU – D]** He complained with a very disappointed look on his face: 'That's so unfair! I'm sure the plane will be delayed.' **[VU – I]** He complained, with a very disappointed look on his face, that this was so unfair because he was sure that the plane would be delayed.

13	Britney is a student at the University of Glasgow. After a heavy snow in the afternoon, she was complaining to her boyfriend James about the weather on their way home. **[VF – D]** Her voice sounded very grumpy and unpleasant: 'I really hate the winter! It's always dark and the roads are too slippery.' **[VF – I]** Her voice sounded very grumpy and unpleasant and unpleasant, saying that she really hated the winter because it's always dark and the roads are too slippery. **[VU – D]** She said, with an unpleasant look on her face: 'I really hate the winter! It's always dark and the roads are too slippery.' **[VU – I]** She said, with an unpleasant look on her face, that she really hated the winter because it's always dark and the roads are too slippery.	James is a student at the University of Glasgow. After a heavy snow in the afternoon, he was complaining to his girlfriend Britney about the weather on their way home. **[VF – D]** His voice sounded very grumpy and unpleasant: 'I really hate the winter! It's always dark and the roads are too slippery.' **[VF – I]** His voice sounded very grumpy and unpleasant, saying that he really hated the winter because it's always dark and the roads are too slippery. **[VU – D]** He said, with an unpleasant look on his face: 'I really hate the winter! It's always dark and the roads are too slippery.' **[VU – I]** He said, with an unpleasant look on his face, that he really hated the winter because it's always dark and the roads are too slippery.
14	Jasmine had been ill for a week so her brother Colin took her to the doctor. After an hour, Jasmine came back to the waiting room looking rather disgruntled. **[VF – D]** She complained in a sharp and irate voice: 'The treatment was absolutely useless! Next time, I'd rather go to a specialist.' **[VF – I]** She complained in a sharp and irate voice that the treatment was absolutely useless and that, next time, she'd rather go to a specialist. **[VU – D]** She complained with an annoyed look on her face: 'The treatment was absolutely useless! Next time, I'd rather go to a specialist.' **[VU – I]** She complained, with an annoyed look on her face, that the treatment was absolutely useless and that, next time, she'd rather go to a specialist.	Sam had been ill for a week so his sister Jasmine took him to the doctor. After an hour, Sam came back to the waiting room looking rather disgruntled. **[VF – D]** He complained in a sharp and irate voice: 'The treatment was absolutely useless! Next time, I'd rather go to a specialist.' **[VF – I]** He complained in a sharp and irate voice that the treatment was absolutely useless and that, next time, he'd rather go to a specialist. **[VU – D]** He complained with an annoyed look on his face: 'The treatment was absolutely useless! Next time, I'd rather go to a specialist.' **[VU – I]** He complained, with an annoyed look on his face, that the treatment was absolutely useless and that, next time, he'd rather go to a specialist.
15	A famous symphony was going to be played for free at the university concert hall. Kate and Andrew were reading the advert, and Kate looked particularly interested. **[VF – D]** Her voice went cheerful and excited: 'It sounds really good and it's for free! I'm definitely going to that concert.' **[VF – I]** Her voice went cheerful and excited, saying that it sounded really good and that it was for free, adding that she would definitely go to that concert. **[VU – D]** With her eyes wide open, she said: 'It sounds really good and it's for free! I'm definitely going to that concert.' **[VU – I]** With her eyes wide open, she said that it sounded really good and that it was for free, adding that she would definitely go to that concert.	A famous symphony was going to be played for free at the university concert hall. Andrew and Kate were reading the advert, and Andrew looked particularly interested. **[VF – D]** His voice went cheerful and excited: 'It sounds really good and it's for free! I'm definitely going to that concert.' **[VF – I]** His voice went cheerful and excited, saying that it sounded really good and that it was for free, adding that he would definitely go to that concert. **[VU – D]** With his eyes wide open, he said: 'It sounds really good and it's for free! I'm definitely going to that concert.' **[VU – I]** With his eyes wide open, he said that it sounded really good and that it was for free, adding that he would definitely go to that concert.

16	On Valentine's Day, Simon at long last worked up the courage to propose to Emma. Emma wasn't sure, but at the same time, she didn't want to upset him. **[VF – D]** So she explained with a soft and gentle voice: 'You are a really nice bloke, Simon, but I'm simply not ready for such a commitment.' **[VF – I]** So she explained with a soft and gentle voice that although she thought he was a really nice bloke, she was simply not ready for such a commitment. **[VU – D]** So she returned an apologetic smile, explaining: 'You are a really nice bloke, Simon, but I'm simply not ready for such a commitment.' **[VU – I]** So she returned an apologetic smile, explaining that although she thought he was a really nice bloke, she was simply not ready for such a commitment.	On Valentine's Day, Emma at long last worked up the courage to express her affection for Simon. Simon wasn't sure, but at the same time, he didn't want to upset her. **[VF – D]** So he explained with a soft and gentle voice: 'You are a really nice girl, Emma, but I'm simply not ready for a relationship.' **[VF – I]** So he explained with a soft and gentle voice that although he thought she was a really nice girl, he was simply not ready for a relationship. **[VU – D]** So he returned an apologetic smile, explaining: 'You are a really nice girl, Emma, but I'm simply not ready for a relationship.' **[VU – I]** So he returned an apologetic smile, explaining that although he thought she was a really nice girl, he was simply not ready for a relationship.
17	Two college students were talking about on-line shopping. Fraser thought it was pointless and boring, but Brenda couldn't disagree more. **[VF – D]** She raised her voice in disagreement: 'What!? On-line shopping is absolutely amazing! I just snapped up some really cheap DVDs on Amazon.' **[VF – I]** She raised her voice in disagreement, saying that she found on-line shopping absolutely amazing, and that she had just snapped up some really cheap DVDs on Amazon. **[VU – D]** She said, gesturing frantically: 'What!? On-line shopping is absolutely amazing! I just snapped up some really cheap DVDs on Amazon.' **[VU – I]** She said, gesturing frantically, that she found on-line shopping absolutely amazing, and that she had just snapped up some really cheap DVDs on Amazon.	Two college students were talking about on-line shopping. Brenda thought it was pointless and boring, but Fraser couldn't disagree more. **[VF – D]** He raised his voice in disagreement: 'What!? On-line shopping is absolutely amazing! I just snapped up some really cheap DVDs on Amazon.' **[VF – I]** He raised his voice in disagreement, saying that he found on-line shopping absolutely amazing, and that he had just snapped up some really cheap DVDs on Amazon. **[VU – D]** He said, gesturing frantically: 'What!? On-line shopping is absolutely amazing! I just snapped up some really cheap DVDs on Amazon.' **[VU – I]** He said, gesturing frantically, that he found on-line shopping absolutely amazing, and that he had just snapped up some really cheap DVDs on Amazon.
18	It was quarter past seven and the Aberdeen Stock House was serving dinner. Debbie and Mike were very pleased with the meal and Debbie in particular was totally smitten. **[VF – D]** Her voice was honeyed and passionate: 'Darling, we have to come here more often. The fish was really tasty and the soup was absolutely delicious.' **[VF – I]** Her voice was honeyed and passionate, saying that they should go there more often because the fish was really fresh and the soup was absolutely delicious. **[VU – D]** There was a pleased look on her face while she said: 'Darling, we have to come here more often. The fish was really tasty and the soup was absolutely delicious.' **[VU – I]** There was a pleased look on her face while she said that they should go there more often because the fish was really fresh and the soup was absolutely delicious.	It was quarter past seven and the Aberdeen Stock House was serving dinner. Mike and Debbie were very pleased with the meal and Mike in particular was totally smitten. **[VF – D]** His voice was full of delight: 'Darling, we have to come here more often. The fish was really tasty and the soup was absolutely delicious.' **[VF – I]** His voice was full of delight, saying that they should go there more often because the fish was really fresh and the soup was absolutely delicious. **[VU – D]** There was a pleased look on his face while he said: 'Darling, we have to come here more often. The fish was really tasty and the soup was absolutely delicious.' **[VU – I]** There was a pleased look on his face while he said that they should go there more often because the fish was really fresh and the soup was absolutely delicious.

19	It was 6-year-old Herbert's first day at school. His mother Laura, who had never let him play with other children before, was very anxious. **[VF – D]** She was talking to her husband with an apprehensive tone in her voice: 'I'm really not sure whether Herbert is ready for this. I'd rather wait for another year.' **[VF – I]** She was talking to her husband, with an apprehensive tone in her voice, saying that she was really not sure whether Herbert was ready for this, and that she'd rather wait for another year. **[VU – D]** She said to her husband, with an apprehensive look on her face: 'I'm really not sure whether Herbert is ready for this. I'd rather wait for another year.' **[VU – I]** She said to her husband, with an apprehensive look on her face, that she was really not sure whether Herbert was ready for this, and that she'd rather wait for another year.	It was 6-year-old Laura's first day at school. His father Herbert, who had never let her play with other children before, was very anxious. **[VF – D]** He was talking to his wife with an apprehensive tone in his voice: 'I'm really not sure whether Laura is ready for this. I'd rather wait for another year.' **[VF – I]** He was talking to his wife, with an apprehensive tone in his voice, saying that he was really not sure whether Laura was ready for this, and that he'd rather wait for another year. **[VU – D]** He said to his wife, with an apprehensive look on his face: 'I'm really not sure whether Laura is ready for this. I'd rather wait for another year.' **[VU – I]** He said to his wife, with an apprehensive look on his face, that he was really not sure whether Laura was ready for this, and that he'd rather wait for another year.
20	In order to finish the project in time, Audrey had been working in the office from 7am to 8pm without taking a break. Her colleague Sean just came back from dinner with a big smile on his face, which obviously upset her a bit. **[VF – D]** She complained, with her voice turned screechy: 'It's so unfair! My stomach had been rumbling all day. I could eat a whole elephant.' **[VF – I]** She complained, with her voice turned screechy, that it was so unfair because her stomach had been rumbling all day, and that she could eat a whole elephant. **[VU – D]** She complained, with an expression of discontent: 'It's so unfair! My stomach had been rumbling all day. I could eat a whole elephant.' **[VU – I]** She complained, with an expression of discontent, that it was so unfair because her stomach had been rumbling all day, and that she could eat a whole elephant.	In order to finish the project in time, Sean had been working in the office from 7am to 8pm without taking a break. His colleague Audrey just came back from dinner with a big smile on her face, which obviously upset him a bit. **[VF – D]** He complained, with his voice turned screechy: 'It's so unfair! My stomach had been rumbling all day. I could eat a whole elephant.' **[VF – I]** He complained, with his voice turned screechy, that it was so unfair because his stomach had been rumbling all day, and that he could eat a whole elephant. **[VU – D]** He complained, with an expression of discontent: 'It's so unfair! My stomach had been rumbling all day. I could eat a whole elephant.' **[VU – I]** He complained, with an expression of discontent, that it was so unfair because his stomach had been rumbling all day, and that he could eat a whole elephant.
21	Cheryl and Barry were in the waiting lounge for their honeymoon flight to Paris. Shortly before boarding time, Cheryl noticed that her handbag was damaged. **[VF – D]** She was shocked and bursted into a panicking shout: 'Oh no! There's a hole in my bag! Don't tell me I've lost my passport!...' **[VF – I]** She was shocked and bursted into a panicking shout that there was a hole in her bag. She was worried that she had lost her passport. **[VU –D]** She was grabbing her hair, obviously shocked: 'Oh no! There's a hole in my bag! Don't tell me I've lost my passport!...' **[VU – I]** She was grabbing her hair, obviously shocked, saying that there was a hole in her bag. She was worried that she had lost her passport.	Barry and Cheryl were in the waiting lounge for their honeymoon flight to Paris. Shortly before boarding time, Barry noticed that his briefcase was damaged. **[VF – D]** He was shocked and bursted into a panicking shout: 'Oh no! There's a hole in my briefcase! Don't tell me I've lost my passport!...' **[VF – I]** He was shocked and bursted into a panicking shout that there was a hole in his briefcase. He was worried that he had lost his passport. **[VU – D]** He was grabbing his hair, obviously shocked: 'Oh no! There's a hole in my briefcase! Don't tell me I've lost my passport!...' **[VU – I]** He was grabbing his hair, obviously shocked, saying that there was a hole in his briefcase. He was worried that he had lost his passport.

22	It was eight o' clock in the evening when the Britain's Got Talent live recording session had finished. On their way out of the theatre, Sarah and Ross were having a giggle over some of the auditions. **[VF – D]** Still chuckling, Sarah said: 'The jugglers at the end were funny, weren't they? Quite pathetic, actually, but very entertaining.' **[VF –I]** Still chuckling, Sarah said that the jugglers at the end were very funny – quite pathetic, actually, but very entertaining. **[VU – D]** Sarah said, with a bemused look on her face: 'The jugglers at the end were funny, weren't they? Quite pathetic, actually, but very entertaining.' **[VU – I]** Sarah said, with a bemused look on her face, that the jugglers at the end were very funny – quite pathetic, actually, but very entertaining.	It was eight o' clock in the evening when the Britain's Got Talent live recording session had finished. On their way out of the theatre, Ross and Sarah were having a giggle over some of the auditions. **[VF – D]** Still chuckling, Ross said: 'The jugglers at the end were funny, weren't they? Quite pathetic, actually, but very entertaining.' **[VF – I]** Still chuckling, Ross said that the jugglers at the end were very funny – quite pathetic, actually, but very entertaining. **[VU – D]** Ross said, with a bemused look on his face: 'The jugglers at the end were funny, weren't they? Quite pathetic, actually, but very entertaining.' **[VU – I]** Ross said, with a bemused look on his face, that the jugglers at the end were very funny – quite pathetic, actually, but very entertaining.
23	Jessica and John were enjoying a hot and sunny day at the holiday resort in Spain. When they were heading towards their usual spot on the beach, Jessica was slightly overwhelmed by the heat. **[VF – D]** She sounded exhausted when she said: 'Phew, this is almost a bit too hot for me! Perhaps we should first go back for a drink under the sunshade.' **[VF – I]** She sounded exhausted when she said that it was almost a bit too hot for her, and suggested to first go back for a drink under the sunshade. **[VU – D]** She said, wiping her forehead with a towel: 'Phew, this is almost a bit too hot for me! Perhaps we should first go back for a drink under the sunshade.' **[VU – I]** She said, wiping her forehead with a towel, that it was almost a bit too hot for her, and suggested to first go back for a drink under the sunshade.	John and Jessica were enjoying a hot and sunny day at the holiday resort in Spain.When they were heading towards their usual spot on the beach, John was slightly overwhelmed by the heat. **[VF – D]** He sounded exhausted when he said: 'Phew, this is almost a bit too hot for me! Perhaps we should first go back for a drink under the sunshade.' **[VF – I]** He sounded exhausted when he said that it was almost a bit too hot for him, and suggested to first go back for a drink under the sunshade. **[VU – D]** He said, wiping his forehead with a towel: 'Phew, this is almost a bit too hot for me! Perhaps we should first go back for a drink under the sunshade.' **[VU – I]** He said, wiping his forehead with a towel, that it was almost a bit too hot for him, and suggested to first go back for a drink under the sunshade.
24	In the small town of Oban lies this charming little pub. One day, famous writer Aileen burst in with a big smile on her face. **[VF – D]** Her voice burst with joy: 'Guess what – my new novel has been accepted for publication! I'll buy everyone a round.' **[VF – I]** Her voice burst with joy, saying that her new novel had been accepted for publication and that she wanted to buy everyone a round. **[VU – D]** She said, waving her arms: 'Guess what – my new novel has been accepted for publication! I'll buy everyone a round.' **[VU – I]** She said, waving her arms, that her new novel had been accepted for publication and that she wanted to buy everyone a round.	In the small town of Oban lies this charming little pub.One day, famous writer Alan burst in with a big smile on his face. **[VF – D]** His voice burst with joy: 'Guess what – my new novel has been accepted for publication! I'll buy everyone a round.' **[VF – I]** His voice burst with joy, saying that his new novel had been accepted for publication and that he wanted to buy everyone a round. **[VU – D]** He said, waving his arms: 'Guess what – my new novel has been accepted for publication! I'll buy everyone a round.' **[VU – I]** He said, waving his arms, that his new novel had been accepted for publication and that he wanted to buy everyone a round.

25	Theatrical agent Peter was visited by one of his clients. Shauna, a talented but struggling actress and singer, was desperate for a part in the new musical. **[VF – D]** She explained, with a hint of anxiety in her voice: 'It's embarrassing to say, but without this job, I won't even be able to pay for electricity.' **[VF – I]** She explained, with a hint of anxiety in her voice, that it was embarrassing to say that without this job, she wouldn't even be able to pay for electricity. **[VU – D]** She explained, with her head swivelled round to avoid eye contact: 'It's embarrassing to say, but without this job, I won't even be able to pay for electricity.' **[VU – I]** She explained, with her head swivelled round to avoid eye contact, that it was embarrassing to say that without this job, she wouldn't even be able to pay for electricity.	Theatrical agent Shauna was visited by one of her clients. Peter, a talented but struggling young actor and singer, was desperate for a part in the new musical. **[VF – D]** He explained, with a hint of anxiety in his voice: 'It's embarrassing to say, but without this job, I won't even be able to pay for electricity.' **[VF – I]** He explained, with a hint of anxiety in his voice, that it was embarrassing to say that without this job, he wouldn't even be able to pay for electricity. **[VU – D]** He explained, with his head swivelled round to avoid eye contact: 'It's embarrassing to say, but without this job, I won't even be able to pay for electricity.' **[VU – I]** He explained, with his head swivelled round to avoid eye contact, that it was embarrassing to say that without this job, he wouldn't even be able to pay for electricity.
26	It was Sunday afternoon when housewife Heather was tidying up the bedroom. After just two minutes, she came out of the bedroom and confronted her husband Ben. **[VF – D]** She was shouting in a raspy voice: 'I can't believe how messy you are! The first thing I found on the floor was your dirty pyjamas!' **[VF – I]** She was shouting in a raspy voice that she could not believe how messy he was, and that the first thing she found on the floor was his dirty pyjamas. **[VU – D]** She said, pointing at the bedroom door: 'I can't believe how messy you are! The first thing I found on the floor was your dirty pyjamas!' **[VU – I]** She said, pointing at the bedroom door, that she could not believe how messy he was, and that the first thing she found on the floor was his dirty pyjamas.	It was Sunday afternoon when father Ben was tidying up his daughter Heather's bedroom. After just two minutes, he came out of the bedroom and confronted Heather. **[VF – D]** Ben was shouting in a raspy voice: 'I can't believe how messy you are! The first thing I found on the floor was your dirty pyjamas!' **[VF – I]** Ben was shouting in a raspy voice that he could not believe how messy she was, and that the first thing he found on the floor was her dirty pyjamas. **[VU – D]** He said, pointing at the bedroom door: 'I can't believe how messy you are! The first thing I found on the floor was your dirty pyjamas!' **[VU – I]** He said, pointing at the bedroom door, that he could not believe how messy she was, and that the first thing he found on the floor was her dirty pyjamas.
27	Morton needed to speak to his project supervisor Marianne at her office. When he asked whether she could spare a minute, she seemed rather busy. **[VF – D]** She replied in a gentle and apologetic tone: 'I'm really sorry, but I have to first finish marking some essays.' **[VF – I]** She replied in a gentle and apologetic tone that she was really sorry and that she had to first finish marking some essays. **[VU – D]** She replied, with an apologetic look on her face: 'I'm really sorry, but I have to first finish marking some essays.' **[VU – I]** She replied, with an apologetic look on her face, that she was really sorry and that she had to first finish marking some essays.	Marianne needed to speak to her project supervisor Morton at his office. When she asked whether he could spare a minute, he seemed rather busy. **[VF – D]** He replied in a gentle and apologetic tone: 'I'm really sorry, but I have to first finish marking some essays.' **[VF – I]** He replied in a gentle and apologetic tone that he was really sorry and that he had to first finish marking some essays. **[VU – D]** He replied, with an apologetic look on his face: 'I'm really sorry, but I have to first finish marking some essays.' **[VU – I]** He replied, with an apologetic look on his face, that he was really sorry and that he had to first finish marking some essays.

28	It was the Christmas Eve and everyone was drinking and dancing at the local pub. Judith was already quite tipsy when she noticed that George was with a girl she had not seen before. **[VF – D]** She went over, and in a playful voice she said: 'George, you've actually never told me that you had such a beautiful sister.' **[VF – I]** She went over, and in a playful voice she said that he actually never told her that he had such a beautiful sister. **[VU – D]** She went over, smiled, and said: 'George, you've actually never told me that you had such a beautiful sister.' **[VU – I]** She went over, smiled, and said that he actually never told her that he had such a beautiful sister.	It was the Christmas Eve and everyone was drinking and dancing at the local pub. George was already quite tipsy when he noticed that Judith was with a girl he had not seen before. **[VF – D]** He went over, and in a playful voice he said: 'Jude, you've actually never told me that you had such a beautiful sister.' **[VF – I]** He went over, and in a playful voice he said that she actually never told him that she had such a beautiful sister. **[VU – D]** He went over, smiled, and said: 'Jude, you've actually never told me that you had such a beautiful sister.' **[VU – I]** He went over, smiled, and said that she actually never told him that she had such a beautiful sister.
29	Maureen had helped David a lot with his coursework, so David wanted to invite her for dinner. Asked what kind of food she would fancy, Maureen first thought for a second. **[VF – D]** Then she replied, with some hesitation in her voice: 'I'm normally not that adventurous, but perhaps I should try some sushi.' **[VF – I]** Then she replied, with some hesitation in her voice, that although she would normally not be that adventurous, perhaps she should try some sushi. **[VU – D]** Then she replied, with an indecisive look on her face: 'I'm normally not that adventurous, but perhaps I should try some sushi.' **[VU – I]** Then she replied, with an indecisive look on her face, that although she would normally not be that adventurous, perhaps she should try some sushi.	David had helped Maureen a lot with her coursework, so Maureen wanted to invite him for dinner. Asked what kind of food he would fancy, David first thought for a second. **[VF – D]** Then he replied, with some hesitation in his voice: 'I'm normally not that adventurous, but perhaps I should try some sushi.' **[VF – I]** Then he replied, with some hesitation in his voice, that although he would normally not be that adventurous, perhaps he should try some sushi. **[VU – D]** Then he replied, with an indecisive look on his face: 'I'm normally not that adventurous, but perhaps I should try some sushi.' **[VU – I]** Then he replied, with an indecisive look on his face, that although he would normally not be that adventurous, perhaps he should try some sushi.
30	Jane was a new post-graduate at the department. When she signed up at the Sports Club, she was having a conversation with Gregg about her favourite sports. **[VF – D]** She said, not without pride in her voice: 'Oh, I haven't practised a lot recently, but I used to be quite good at tennis.' **[VF – I]** She said, not without pride in her voice, that although she had not practised a lot recently, she used to be quite good at tennis. **[VU – D]** She said, mocking a tennis serve: 'Oh, I haven't practised a lot recently, but I used to be quite good at tennis.' **[VU – I]** She said, mocking a tennis serve, that although she had not practised a lot recently, she used to be quite good at tennis.	Gregg was a new post-graduate at the department. When he signed up at the Sports Club, he was having a conversation with Jane about his favourite sports. **[VF – D]** He said, not without pride in his voice: 'Oh, I haven't practised a lot recently, but I used to be quite good at tennis.' **[VF – I]** He said, not without pride in his voice, that although he had not practised a lot recently, he used to be quite good at tennis. **[VU – D]** He said, mocking a tennis serve: 'Oh, I haven't practised a lot recently, but I used to be quite good at tennis.' **[VU – I]** He said, mocking a tennis serve, that although he had not practised a lot recently, he used to be quite good at tennis.

31	Betty and Leon were trying to get a mortgage for a new flat. While Leon was studying the relevant newspaper adverts, Betty was browsing some web-sites. **[VF – D]** After a while, she sounded quite disillusioned: 'Hmm, these days, it looks as if nobody would give us an instant cashback.' **[VF – I]** After a while, she sounded quite disillusioned, saying that it looked as if nobody would give them an instant cashback. **[VU – D]** After a while, she sipped her tea and said: 'Hmm, these days, it looks as if nobody would give us an instant cashback.' **[VU – I]** After a while, she sipped her tea and said that it looked as if nobody would give them an instant cashback.	Leon and Betty were trying to get a mortgage for a new flat. While Betty was studying the relevant newspaper adverts, Leon was browsing some web-sites. **[VF – D]** After a while, he sounded quite disillusioned: 'Hmm, these days, it looks as if nobody would give us an instant cashback.' **[VF – I]** After a while, he sounded quite disillusioned, saying that it looked as if nobody would give them an instant cashback. **[VU – D]** After a while, he sipped his tea and said: 'Hmm, these days, it looks as if nobody would give us an instant cashback.' **[VU – I]** After a while, he sipped his tea and said that it looked as if nobody would give them an instant cashback.
32	Helen and Marcus were planning a trip to Europe for their wedding anniversary. Helen was leaning more towards Prague although Marcus preferred Rome. **[VF – D]** Helen said in a cheeky voice: 'Okay, Rome is nice, but the Czechs actually have much better beer.' **[VF – I]** Helen said in a cheeky voice that Rome would be nice, but that the Czechs would actually have much better beer. **[VU – D]** Helen explained, with a cheeky smile: 'Okay, Rome is nice, but the Czechs actually have much better beer.' **[VU – I]** Helen explained, with a cheeky smile, that Rome would be nice, but that the Czechs would actually have much better beer.	Marcus and Helen were planning a trip to Europe for their wedding anniversary. Marcus was leaning more towards Prague although Helen preferred Rome. **[VF – D]** Marcus said in a cheeky voice: 'Okay, Rome is nice, but the Czechs actually have much better beer.' **[VF – I]** Marcus said in a cheeky voice that Rome would be nice, but that the Czechs would actually have much better beer. **[VU – D]** Marcus explained, with a cheeky smile: 'Okay, Rome is nice, but the Czechs actually have much better beer.' **[VU – I]** Marcus explained, with a cheeky smile, that Rome would be nice, but that the Czechs would actually have much better beer.
33	Dianne and Bill went out bowling together. When Bill wanted to pick a ball of his favourite colour, Dianne looked a bit worried. **[VF – D]** She said, in a caring and considerate tone: 'That one looks quite filthy. I'd rather pick the green one which I think is much cleaner.' **[VF – I]** She said, in a caring and considerate tone, that the ball looked quite filthy and that she'd rather pick the green one which she thought was much cleaner. **[VU – D]** She said, with a frown on her face: 'That one looks quite filthy. I'd rather pick the green one which I think is much cleaner.' **[VU – I]** She said, with a frown on her face, that the ball looked quite filthy and that she'd rather pick the green one which she thought was much cleaner.	Bill and Dianne went out bowling together. When Dianne wanted to pick a ball of her favourite colour, Bill looked a bit worried. **[VF – D]** He said, in a caring and considerate tone: 'That one looks quite heavy. I'd rather pick the green one which I think is much smaller.' **[VF – I]** He said, in a caring and considerate tone, that the ball looked quite heavy and that he'd rather pick the green one which he thought was much smaller. **[VU – D]** He said, with a frown on his face: 'That one looks quite heavy. I'd rather pick the green one which I think is much smaller.' **[VU – I]** He said, with a frown on his face, that the ball looked quite heavy and that he'd rather pick the green one which he thought was much smaller.

34	Becky and Rob were buying duvet covers at the store. When Rob chose a bright red one, Becky was completely against the colour. **[VF – D]** In a piercing, sarcastic tone, she said: 'No way! Just imagine that every morning you'd be waking up in a sea of blood!' **[VF – I]** In a piercing, sarcastic tone, she asked Rob just to imagine that every morning he would be waking up in a sea of blood. **[VU – D]** She said, with her head shaking frantically: 'No way! Just imagine that every morning you'd be waking up in a sea of blood!' **[VU – I]** She said, with her head shaking frantically, just to imagine that every morning he would be waking up in a sea of blood.	Rob and Becky were buying duvet covers at the store. When Becky chose a bright red one, Rob was completely against the colour. **[VF – D]** In a piercing, sarcastic tone, he said: 'No way! Just imagine that every morning you'd be waking up in a sea of blood!' **[VF – I]** In a piercing, sarcastic tone, he asked Becky just to imagine that every morning she would be waking up in a sea of blood. **[VU – D]** He said, with his head shaking frantically: 'No way! Just imagine that every morning you'd be waking up in a sea of blood!' **[VU – I]** He said, with his head shaking frantically, just to imagine that every morning she would be waking up in a sea of blood.
35	Leona and Harry were enjoying an afternoon walk in the park. When they were approaching the pond, Leona was impressed by this peaceful place. **[VF – D]** Her voice sounded totally smitten and cheerful: 'This is so relaxing after a busy week! It is almost magical.' **[VF – I]** Her voice sounded totally smitten and cheerful, saying that it was so relaxing after a busy week, and that it was almost magical. **[VU – D]** She closed her eyes and took a deep breath: 'This is so relaxing after a busy week! It is almost magical.' **[VU – I]** She closed her eyes and took a deep breath, saying that it was so relaxing after a busy week, and that it was almost magical.	Harry and Leona were enjoying an afternoon walk in the park. When they were approaching the pond, Harry was impressed by this peaceful place. **[VF – D]** His voice sounded totally smitten and cheerful: 'This is so relaxing after a busy week! It is almost magical.' **[VF – I]** His voice sounded totally smitten and cheerful, saying that it was so relaxing after a busy week, and that it was almost magical. **[VU – D]** He closed his eyes and took a deep breath: 'This is so relaxing after a busy week! It is almost magical.' **[VU – I]** He closed his eyes and took a deep breath, saying that it was so relaxing after a busy week, and that it was almost magical.
36	Schoolteacher Lynda recently caused a stir by marrying Edward, one of her ex-pupils. Annoyed by all the gossip, Lynda told Edward to simply not listen to what people say. **[VF – D]** She said, in a firm voice: 'My dear, I think there's nothing wrong about a teacher marrying a former student.' **[VF – I]** She said, in a firm voice, that she thought there was nothing wrong about a teacher marrying a former student. **[VU – D]** With a resolute expression on her face, she said: 'My dear, I think there's nothing wrong about a teacher marrying a former student.' **[VU – I]** With a resolute expression on her face, she said that she thought there was nothing wrong about a teacher marrying a former student.	Schoolteacher Edward recently caused a stir by marrying Lynda, one of his ex-pupils. Annoyed by all the gossip, Edward told Lynda to simply not listen to what people say. **[VF – D]** He said, in a firm voice: 'My dear, I think there's nothing wrong about a teacher marrying a former student.' **[VF – I]** He said, in a firm voice, that he thought there was nothing wrong about a teacher marrying a former student. **[VU – D]** With a resolute expression on his face, he said: 'My dear, I think there's nothing wrong about a teacher marrying a former student.' **[VU – I]** With a resolute expression on his face, he said that he thought there was nothing wrong about a teacher marrying a former student.

37	After making a small fortune in the lottery, Ralph bought quite an expensive car. Today in the car park, his colleague Pamela was extremely impressed by his new possession. **[VF – D]** Her voice wavered between admiration and envy: 'Wow! Is that your new car? It looks really stylish.' **[VF – I]** Her voice wavered between admiration and envy when she asked him whether that was his new car, adding that it looked really stylish. **[VU – D]** Her face wavered between admiration and envy: 'Wow! Is that your new car? It looks really stylish.' **[VU – I]** Her face wavered between admiration and envy when she asked him whether that was his new car, adding that it looked really stylish.	After making a small fortune in the lottery, Pamela bought quite an expensive car. Today in the car park, her colleague Ralph was extremely impressed by her new possession. **[VF – D]** His voice wavered between admiration and envy: 'Wow! Is that your new car? It looks really stylish.' **[VF – I]** His voice wavered between admiration and envy when he asked her whether that was her new car, adding that it looked really stylish. **[VU – D]** His face wavered between admiration and envy: 'Wow! Is that your new car? It looks really stylish.' **[VU – I]** His face wavered between admiration and envy when he asked her whether that was her new car, adding that it looked really stylish.
38	Katy invited some of her friends home and started making a mess in the kitchen. Her mum was obviously not happy about this, and Katy was quick to offer an excuse. **[VF – D]** Her voice sounded apprehensive: 'Well, we just had a couple of drinks because the pubs are already closed.' **[VF – I]** Her voice sounded apprehensive, saying that they just had a couple of drinks because the pubs were already closed. **[VU – D]** She was staring at the floor: 'Well, we just had a couple of drinks because the pubs are already closed.' **[VU – I]** She was staring at the floor, saying that they just had a couple of drinks because the pubs were already closed.	Brian invited some of his of mates home and started making a mess in the kitchen. His mum was obviously not happy about this, and Brian was quick to offer an excuse. **[VF – D]** His voice sounded apprehensive: 'Well, we just had a couple of drinks because the pubs are already closed.' **[VF – I]** His voice sounded apprehensive, saying that they just had a couple of drinks because the pubs were already closed. **[VU – D]** He was staring at the floor: 'Well, we just had a couple of drinks because the pubs are already closed.' **[VU – I]** He was staring at the floor, saying that they just had a couple of drinks because the pubs were already closed.
39	Maddie was injured in a car crash last week. At the hospital, she was quickly recovering. Today, she was surprised that some distant relatives came to visit her. **[VF – D]** Maddie sounded sarcastic: 'Oh, looks as if one has to get hurt to get some flowers.' **[VF – I]** Maddie sounded sarcastic, saying that it looked as if one had to get hurt to get some flowers. **[VU – D]** Maddie said, with a sarcastic smile on her face: 'Oh, looks as if one has to get hurt to get some flowers.' **[VU – I]** Maddie said, with a sarcastic smile on her face, that it looked as if one had to get hurt to get some flowers.	Daniel was injured in a car crash last week. At the hospital, he was quickly recovering. Today, he was surprised that some distant relatives came to visit him. **[VF – D]** Daniel sounded sarcastic: 'Oh, looks as if one has to get hurt to get some flowers.' **[VF – I]** Daniel sounded sarcastic, saying that it looked as if one had to get hurt to get some flowers. **[VU – D]** Daniel said, with a sarcastic smile on his face: 'Oh, looks as if one has to get hurt to get some flowers.' **[VU – I]** Daniel said, with a sarcastic smile on his face, that it looked as if one had to get hurt to get some flowers.

40	Cindy had promised Jacob to buy a bonsai tree to decorate their living room. When Jacob wanted to place the plant by the television, she tried to persuade him of an alternative option. **[VF – D]** She explained in a clear and firm voice: 'Look, if you do that, there will be no sunlight for the bonsai, so it's better to place it by the window.' **[VF – I]** She explained in a clear and firm voice that if he did that, there would be no sunlight for the bonsai, so it was better to place it by the window. **[VU – D]** She explained, with a wagging forefinger: 'Look, if you do that, there will be no sunlight for the bonsai, so it's better to place it by the window.' **[VU – I]** She explained, with a wagging forefinger, that if he did that, there would be no sunlight for the bonsai, so it was better to place it by the window.	Jacob had promised Cindy to buy a bonsai tree to decorate their living room. When Cindy wanted to place the plant by the television, he tried to persuade her of an alternative option. **[VF – D]** He explained in a clear and firm voice: 'Look, if you do that, there will be no sunlight for the bonsai, so it's better to place it by the window.' **[VF – I]** He explained in a clear and firm voice that if she did that, there would be no sunlight for the bonsai, so it was better to place it by the window. **[VU – D]** He explained, with a wagging forefinger: 'Look, if you do that, there will be no sunlight for the bonsai, so it's better to place it by the window.' **[VU – I]** He explained, with a wagging forefinger, that if she did that, there would be no sunlight for the bonsai, so it was better to place it by the window.
41	PhD student Jim was summoned to her supervisor Ella's office to give a report on his current progress. When Jim asked for an extension, Ella looked concerned. **[VF – D]** Her voice turned serious: 'Hmm, we really need those data in by next month for that conference.' **[VF – I]** Her voice turned serious, saying that they really needed those data in by next month for that conference. **[VU – D]** Her face turned serious: 'Hmm, we really need those data in by next month for that conference.' **[VU – I]** Her face turned serious, saying that they really needed those data in by next month for that conference.	PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. When Ella asked for an extension, Jim looked concerned. **[VF – D]** His voice turned serious: 'Hmm, we really need those data in by next month for that conference.' **[VF – I]** His voice turned serious, saying that they really needed those data in by next month for that conference. **[VU – D]** His face turned serious: 'Hmm, we really need those data in by next month for that conference.' **[VU – I]** His face turned serious, saying that they really needed those data in by next month for that conference.
42	Joanne was taking Albert to the new Schwarzenegger movie. Joanne was very excited, but since Albert seemed not very keen, she tried to persuade him to go with her. **[VF – D]** She raised her voice, sounding very agitated: 'Come on! I'm sure you will love it! The movie has some really cool special effects.' **[VF – I]** She raised her voice, sounding very agitated, and said that she was sure he would love it, and that the movie had some really cool special effects. **[VU – D]** She said, with her eyes wide open: 'Come on! I'm sure you will love it! The movie has some really cool special effects.' **[VU – I]** She said, with her eyes wide open, that she was sure he would love it, and that the movie had some really cool special effects.	Albert was taking Joanne to the new Schwarzenegger movie. Albert was very excited, but since Joanne seemed not very keen, he tried to persuade her to go with him. **[VF – D]** He raised his voice, sounding very agitated: 'Come on! I'm sure you will love it! The movie has some really cool special effects.' **[VF – I]** He raised his voice, sounding very agitated, and said that he was sure she would love it, and that the movie had some really cool special effects. **[VU – D]** He said, with his eyes wide open: 'Come on! I'm sure you will love it! The movie has some really cool special effects.' **[VU – I]** He said, with his eyes wide open, that he was sure she would love it, and that the movie had some really cool special effects.
43	Stephanie and Neil were visiting the local distillery. Stephanie was amazed by the range of different Whiskeys produced, and she tried a very expensive Single Malt. **[VF – D]** She coughed and said in a breathy voice: 'Gee, that's a strong one! I'm glad I won't have to drive us home.' **[VF – I]** She coughed and said in a breathy voice that it was really strong and that she was glad she didn't have to drive them home. **[VU – D]** She said, slapping her chest: 'Gee, that's a strong one! I'm glad I won't have to drive us home.' **[VU – I]** She said, slapping her chest, that it was really strong and that she was glad she didn't have to drive them home.	Neil and Stephanie were visiting the local distillery. Neil was amazed by the range of different Whiskeys produced, and he tried a very expensive Single Malt. **[VF – D]** He coughed and said in a breathy voice: 'Gee, that's a strong one! I'm glad I won't have to drive us home.' **[VF – I]** He coughed and said in a breathy voice that it was really strong and that he was glad he didn't have to drive them home. **[VU – D]** He said, slapping his chest: 'Gee, that's a strong one! I'm glad I won't have to drive us home.' **[VU – I]** He said, slapping his chest, that it was really strong and that he was glad he didn't have to drive them home.

44	Molly and her husband Clive went out to the new seafood restaurant for the first time. Molly was extremely anxious about food poisoning and did not want to have anything raw. **[VF – D]** In a firm voice, she said to Clive: 'I don't care what other people might think, but I will order properly cooked oysters.' **[VF – I]** In a firm voice, she said to Clive that, regardless of what other people might think, she would order properly cooked oysters. **[VU – D]** With a serious look on her face, she said to Clive: 'I don't care what other people might think, but I will order properly cooked oysters.' **[VU – I]** With a serious look on her face, she said to Clive that regardless of what other people might think, she would order properly cooked oysters.	Clive and his wife Molly went out to the new seafood restaurant for the first time. Clive was extremely anxious about food poisoning and did not want to have anything raw. **[VF – D]** In a firm voice, he said to Molly: 'I don't care what other people might think, but I will order properly cooked oysters.' **[VF – I]** In a firm voice, he said to Molly that, regardless of what other people might think, he would order properly cooked oysters. **[VU – D]** With a serious look on his face, he said to Molly: 'I don't care what other people might think, but I will order properly cooked oysters.' **[VU – I]** With a serious look on his face, he said to Molly that regardless of what other people might think, he would order properly cooked oysters.
45	Lucy and her friends were watching a movie at the cinema. Lucy wasn't particularly keen on romantic comedies, and she was complaining a lot after the film. **[VF – D]** Her voice sounded harsh and uncontent: 'God, that movie was terrible! I've never been so bored in my life.' **[VF – I]** Her voice sounded harsh and uncontent, saying that the movie was terrible and that she had never been so bored in her life. **[VU – D]** She was rolling her eyes: 'God, that movie was terrible! I've never been so bored in my life.' **[VU – I]** She was rolling her eyes, saying that the movie was terrible and that she had never been so bored in her life.	Luke and his friends were watching a movie at the cinema. Luke wasn't particularly keen on romantic comedies, and he was complaining a lot after the film. **[VF – D]** His voice sounded harsh and uncontent: 'God, that movie was terrible! I've never been so bored in my life.' **[VF – I]** His voice sounded harsh and uncontent, saying that the movie was terrible and that he had never been so bored in his life. **[VU – D]** He was rolling his eyes: 'God, that movie was terrible! I've never been so bored in my life.' **[VU – I]** He was rolling his eyes, saying that the movie was terrible and that he had never been so bored in his life.
46	On the train to London, Thomas was very nervous about his upcoming job interview. When he reached for a cigarette to calm himself down, the girl next to him complained instantly. **[VF – D]** The girl called out, in a stroppy tone: 'Oi! Smoking on public transport is illegal!' **[VF – I]** The girl called out, in a stroppy tone, that smoking on public transport was illegal. **[VU – D]** The girl called out, with an angry look on her face: 'Oi! Smoking on public transport is illegal!' **[VU – I]** The girl called out, with an angry look on her face, that smoking on public transport was illegal.	On the train to London, Thomas was very nervous about his upcoming job interview. When he reached for a cigarette to calm himself down, the middle-aged man next to him complained instantly. **[VF – D]** The man called out, in a stroppy tone: 'Oi! Smoking on public transport is illegal!' **[VF – I]** The man called out, in a stroppy tone, that smoking on public transport was illegal. **[VU – D]** The man called out, with an angry look on his face: 'Oi! Smoking on public transport is illegal!' **[VU – I]** The man called out, with an angry look on his face, that smoking on public transport was illegal.
47	During a coffee break, university teachers Isabel and Robert met in the common room. Isabel was slightly irritated with the mess in the kitchen. **[VF – D]** Her voice was indignant: 'Look at this! Certain people must always leave their dirty mugs on the table.' **[VF – I]** Her voice was indignant, saying that, obviously, certain people must always leave their dirty mugs on the table. **[VU – D]** She said, pointing at the kitchen table: 'Look at this! Certain people must always leave their dirty mugs on the table.' **[VU – I]** She said, pointing at the kitchen table, that obviously, certain people must always leave their dirty mugs on the table.	During a coffee break, university teachers Robert and Isabel met in the common room. Robert was slightly irritated with the mess in the kitchen. **[VF – D]** His voice was indignant: 'Look at this! Certain people must always leave their dirty mugs on the table.' **[VF – I]** His voice was indignant, saying that, obviously, certain people must always leave their dirty mugs on the table. **[VU – D]** He said, pointing at the kitchen table: 'Look at this! Certain people must always leave their dirty mugs on the table.' **[VU – I]** He said, pointing at the kitchen table, that obviously, certain people must always leave their dirty mugs on the table.

48	School teachers Liana and Duncan were arguing about the curriculum. Apparently, Liana was not a big fan of Duncan's modern style of teaching. **[VF – D]** She sounded arrogant and irreverent: 'If my own children were at this school, I'd make sure they would never have to attend your lessons.' **[VF – I]** She sounded arrogant and irreverent, saying that if her own children were at this school, she'd make sure they would never have to attend his lessons. **[VU – D]** She said, with an arrogant look on her face: 'If my own children were at this school, I'd make sure they would never have to attend your lessons.' **[VU – I]** She said, with an arrogant look on her face, that if her own children were at this school, she'd make sure they would never have to attend his lessons.	School teachers Duncan and Liana were arguing about the curriculum. Apparently, Duncan was not a big fan of Liana's modern style of teaching. **[VF – D]** He sounded arrogant and irreverent: 'If my own children were at this school, I'd make sure they would never have to attend your lessons.' **[VF – I]** He sounded arrogant and irreverent, saying that if his own children were at this school, he'd make sure they would never have to attend her lessons. **[VU – D]** He said, with an arrogant look on his face: 'If my own children were at this school, I'd make sure they would never have to attend your lessons.' **[VU – I]** He said, with an arrogant look on his face, that if his own children were at this school, he'd make sure they would never have to attend her lessons.
49	It was getting quite late and most of Scott's friends were about to leave his party. Andrea in particular seemed to have had a great time when she was leaving. **[VF – D]** Her voice was jolly and cheerful: 'Sorry about the crack in that glass. The party was absolutely fantastic.' **[VF – I]** Her voice was jolly and cheerful when she apologised for the crack in that glass, adding that the party was absolutely fantastic. **[VU – D]** She apologised, with a big smile: 'Sorry about the crack in that glass. The party was absolutely fantastic.' **[VU – I]** She apologised, with a big smile, for the crack in that glass, adding that the party was absolutely fantastic.	It was getting quite late and most of Andrea's friends were about to leave her party. Scott in particular seemed to have had a great time when he was leaving. **[VF – D]** His voice was jolly and cheerful: 'Sorry about the crack in that glass. The party was absolutely fantastic.' **[VF – I]** His voice was jolly and cheerful when he apologised for the crack in that glass, adding that the party was absolutely fantastic. **[VU – D]** He apologised, with a big smile: 'Sorry about the crack in that glass. The party was absolutely fantastic.' **[VU – I]** He apologised, with a big smile, for the crack in that glass, adding that the party was absolutely fantastic.
50	Donna always enjoyed a bit of gardening in her spare time. She recently bought a new lawnmower, but she was quite frustrated because it simply refused to work. **[VF – D]** She complained loudly and resentfully: 'Sod it! Things used to be much better in the past. All one can buy these days is rubbish!' **[VF – I]** She complained loudly and resentfully that things used to be much better in the past, and that all one can buy these days was rubbish. **[VU – D]** She complained with a very long face: 'Sod it! Things used to be much better in the past. All one can buy these days is rubbish!' **[VU – I]** She complained with a very long face that things used to be much better in the past, and that all one can buy these days was rubbish.	Doug always enjoyed a bit of gardening in his spare time. He recently bought a new lawnmower, but he was quite frustrated because it simply refused to work. **[VF – D]** He complained loudly and resentfully: 'Sod it! Things used to be much better in the past. All one can buy these days is rubbish!' **[VF – I]** He complained loudly and resentfully that things used to be much better in the past, and that all one can buy these days was rubbish. **[VU – D]** He complained with a very long face: 'Sod it! Things used to be much better in the past. All one can buy these days is rubbish!' **[VU – I]** He complained with a very long face that things used to be much better in the past, and that all one can buy these days was rubbish.

51	After class, Rachel had a serious conversation with one of his pupils. The boy had been accused of bullying and Rachel was giving him a caution. **[VF – D]** Rachel sounded very formal and serious: 'If the accusations are true, you will be in serious trouble. One cannot tolerate such kind of behaviour.' **[VF – I]** Rachel sounded very formal and serious, saying that if the accusations were true, the boy would be in serious trouble, and that one cannot tolerate such kind of behaviour. **[VU – D]** Rachel said, with a serious expression on her face: 'If the accusations are true, you will be in serious trouble. One cannot tolerate such kind of behaviour.' **[VU – I]** Rachel said, with a serious expression on her face, that if the accusations were true, the boy would be in serious trouble, and that one cannot tolerate such kind of behaviour.	After class, Roy had a serious conversation with one of his pupils. The boy had been accused of bullying and Roy was giving him a caution. **[VF – D]** Roy sounded very formal and serious: 'If the accusations are true, you will be in serious trouble. One cannot tolerate such kind of behaviour.' **[VF – I]** Roy sounded very formal and serious, saying that if the accusations were true, the boy would be in serious trouble, and that one cannot tolerate such kind of behaviour. **[VU – D]** Roy said, with a serious expression on his face: 'If the accusations are true, you will be in serious trouble. One cannot tolerate such kind of behaviour.' **[VU – I]** Roy said, with a serious expression on his face, that if the accusations were true, the boy would be in serious trouble, and that one cannot tolerate such kind of behaviour.
52	In the board room, assistant manager Chloe was giving a presentation on how to improve the company's sales figures. When she noticed that her boss was anything but impressed, Chloe used her lack of experience as an excuse. **[VF – D]** Her voice was shy and hesitant: 'Sorry, this is actually my first ever presentation and I'm probably not very persuasive.' **[VF – I]** Her voice was shy and hesitant, saying that this was actually her first ever presentation, and that she was probably not very persuasive. **[VU – D]** She said, looking rather nervous: 'Sorry, this is actually my first ever presentation and I'm probably not very persuasive.' **[VU – I]** She said, looking rather nervous, that this was actually her first ever presentation, and that she was probably not very persuasive.	In the board room, assistant manager Craig was giving a presentation on how to improve the company's sales figures. When he noticed that his boss was anything but impressed, Craig used his lack of experience as an excuse. **[VF – D]** His voice was shy and hesitant: 'Sorry, this is actually my first ever presentation and I'm probably not very persuasive.' **[VF – I]** His voice was shy and hesitant, saying that this was actually his first ever presentation, and that he was probably not very persuasive. **[VU – D]** He said, looking rather nervous: 'Sorry, this is actually my first ever presentation and I'm probably not very persuasive.' **[VU – I]** He said, looking rather nervous, that this was actually his first ever presentation, and that he was probably not very persuasive.
53	Smith was working in a small antiques shop down the local high street. Today, a middle-aged posh lady with thick glasses came into the shop. **[VF – D]** She looked around and said in a nonchalant tone: 'You may be surprised to learn that I'm a world-renowned collector of rare memorabilia of White-eared Pheasant.' **[VF – I]** She looked around and said, in a nonchalant tone, that he might be surprised to learn that she was a world-renowned collector of rare memorabilia of White-eared Pheasant. **[VU – D]** She looked around and said with a slightly absent facial expression: 'You may be surprised to learn that I'm a world-renowned collector of rare memorabilia of White-eared Pheasant.' **[VU – I]** She looked around and said, with a slightly absent facial expression, that he might be surprised to learn that she was a world-renowned collector of rare memorabilia of White-eared Pheasant.	Alice was working in a small antiques shop down the local high street. Today, a weird-looking man with greasy hair and thick glasses came into the shop. **[VF – D]** He looked around and said in a slightly croaky voice: 'You may be surprised to learn that I'm a world-renowned collector of rare memorabilia of White-eared Pheasant.' **[VF – I]** He looked around and said, in a slightly croaky voice, that she might be surprised to learn that he was a world-renowned collector of rare memorabilia of White-eared Pheasant. **[VU – D]** He looked around and said with a slightly absent facial expression: 'You may be surprised to learn that I'm a world-renowned collector of rare memorabilia of White-eared Pheasant.' **[VU – I]** He looked around and said, with a slightly absent facial expression, that she might be surprised to learn that he was a world-renowned collector of rare memorabilia of White-eared Pheasant.

54	Sandra and Ray were on a trip to the countryside for the weekend. Sandra quite liked the outdoors and particularly that rural smell. **[VF – D]** She said in a relaxed and content voice: 'Aah, this is so much better than crowded cities and pollution.' **[VF – I]** She said, in a relaxed and content voice, that it was so much better than crowded cities and pollution. **[VU – D]** She said, with a content smile: 'Aah, this is so much better than crowded cities and pollution.' **[VU – I]** She said, with a content smile, that it was so much better than crowded cities and pollution.	Ray and Sandra were on a trip to the countryside for the weekend. Ray quite liked the outdoors and particularly that rural smell. **[VF – D]** He said in a relaxed and content voice: 'Aah, this is so much better than crowded cities and pollution.' **[VF – I]** He said, in a relaxed and content voice, that it was so much better than crowded cities and pollution. **[VU – D]** He said, with a content smile: 'Aah, this is so much better than crowded cities and pollution.' **[VU – I]** He said, with a content smile, that it was so much better than crowded cities and pollution.
55	Earlier in the afternoon, Janis and Max were having a cigarette in front of the main entrance. Not knowing what to say, Janis started her typical conversation about the weather. **[VF – D]** She tried to speak in a witty tone: 'Gosh, it's quite windy today – but at least it's not raining.' **[VF – I]** She tried to speak in a witty tone, saying that although it was quite windy today, at least it was not raining. **[VU – D]** She said, puffing away her cigarette: 'Gosh, it's quite windy today – but at least it's not raining.' **[VU –I]** She said, puffing away her cigarette, that although it was quite windy today, at least it was not raining.	Earlier in the afternoon, Max and Janis were having a cigarette in front of the main entrance. Not knowing what to say, Max started his typical conversation about the weather. **[VF- D]** He tried to speak in a witty tone: 'Gosh, it's quite windy today – but at least it's not raining.' **[VF – I]** He tried to speak in a witty tone, saying that although it was quite windy today, at least it was not raining. **[VU – D]** He said, puffing away his cigarette: 'Gosh, it's quite windy today – but at least it's not raining.' **[VU – I]** He said, puffing away his cigarette, that although it was quite windy today, at least it was not raining.
56	A journalist was interviewing Amy, the older sister of a famous actor. Amy did not really enjoy the attention she was given and wasn't sure what to say. **[VF – D]** She proclaimed in a slightly annoyed tone: 'I actually don't see my brother very often these days. He is completely devoted to his career!' **[VF – I]** She proclaimed, in a slightly annoyed tone, that she actually did not see her brother very often in those days, and that he was completely devoted to his career. **[VU – D]** She proclaimed, inspecting her fingernails: 'I actually don't see my brother very often these days. He is completely devoted to his career!' **[VU – I]** She proclaimed, inspecting her fingernails, that she actually did not see her brother very often in those days, and that he was completely devoted to his career.	A journalist was interviewing Eric, the older brother of a famous pop diva. Eric did not really enjoy the attention he was given and wasn't sure what to say. **[VF – D]** He proclaimed in a slightly annoyed tone: 'I actually don't see my sister very often these days. She is completely devoted to her career!' **[VF – I]** He proclaimed, in a slightly annoyed tone, that he actually did not see his sister very often in those days, and that she was completely devoted to her career. **[VU – D]** He proclaimed, inspecting his fingernails: 'I actually don't see my sister very often these days. She is completely devoted to her career!' **[VU – I]** He proclaimed, inspecting his fingernails, that he actually did not see his sister very often in those days, and that she was completely devoted to her career.
57	Sheena went to the local McDonalds and incidentally met Thomas, who he knew from former Weight Watchers meetings. Naturally, they were both quite embarrassed, but Sheena finally broke the silence. **[VF – D]** She tried to sound ironic: 'Weight Watchers is rubbish, isn't it? I've tried so many diets but I'm still overweight.' **[VF – I]** She tried to sound ironic, saying that Weight Watchers was rubbish, and that she had tried so many diets but was still overweight. **[VU – D]** She said, with an ironic wink: 'Weight Watchers is rubbish, isn't it? I've tried so many diets but I'm still overweight.' **[VU – I]** She said, with an ironic wink, that Weight Watchers was rubbish, and that she had tried so many diets but was still overweight.	Thomas went to the local McDonalds and incidentally met Sheena, who he knew from former Weight Watchers meetings. Naturally, they were both quite embarrassed, but Thomas finally broke the silence. **[VF – D]** He tried to sound ironic: 'Weight Watchers is rubbish, isn't it? I've tried so many diets but I'm still overweight.' **[VF – I]** He tried to sound ironic, saying that Weight Watchers was rubbish, and that he had tried so many diets but was still overweight. **[VU – D]** He said, with an ironic wink: 'Weight Watchers is rubbish, isn't it? I've tried so many diets but I'm still overweight.' **[VU – I]** He said, with an ironic wink, that Weight Watchers was rubbish, and that he had tried so many diets but was still overweight.

58	A car-boot sale was taking place in the neighbourhood and Catherine was trying to get rid of some of the old records she owned since the 1970s.A trendy young man appeared to be very interested, so Catherine started advertising. **[VF – D]** She said in an enthusiastic and persuasive tone: 'I can give you three albums for a fiver. You won't find any of that stuff on the internet.' **[VF – I]** She said in an enthusiastic and persuasive tone that she could give him three albums for a fiver, adding that he wouldn't find any of that stuff on the internet. **[VU – D]** She said, holding a stack of CDs in her hands: 'I can give you three albums for a fiver. You won't find any of that stuff on the internet.' **[VU – I]** She said, holding a stack of CDs in her hands, that she could give him three albums for a fiver, adding that he wouldn't find any of that stuff on the internet.	A car-boot sale was taking place in the neighbourhood and Kenny was trying to get rid of some of the old records he owned since the 1970s.A trendy young woman appeared to be very interested, so Kenny started advertising. **[VF – D]** He said in an enthusiastic and persuasive tone: 'I can give you three albums for a fiver. You won't find any of that stuff on the internet.' **[VF – I]** He said in an enthusiastic and persuasive tone that he could give her three albums for a fiver, adding that she wouldn't find any of that stuff on the internet. **[VU – D]** He said, holding a stack of CDs in his hands: 'I can give you three albums for a fiver. You won't find any of that stuff on the internet.' **[VU – I]** He said, holding a stack of CDs in his hands, that he could give her three albums for a fiver, adding that she wouldn't find any of that stuff on the internet.
59	A trendy night club had been recently opened near Patricia's flat and it proved instantly popular. However, Patricia often felt disturbed by the noise, and this morning she was complaining to the neighbours about it. **[VF – D]** Her voice was furious and raspy: 'This is unbelievable! Every time I have to work early hours, they are having a bash!' **[VF – I]** Her voice was furious and raspy, saying that every time she has to work early hours, they are having a bash. **[VU – D]** She said, with her face turning red: 'This is unbelievable! Every time I have to work early hours, they are having a bash!' **[VU – I]** She said, with her face turning red, that every time she has to work early hours, they are having a bash.	A trendy night club had been recently opened near Marvin's flat and it proved instantly popular. However, Marvin often felt disturbed by the noise, and this morning he was complaining to the neighbours about it. **[VF – D]** His voice was furious and raspy: 'This is unbelievable! Every time I have to work early hours, they are having a bash!' **[VF – I]** His voice was furious and raspy, saying that every time he has to work early hours, they are having a bash. **[VU – D]** He said, with his face turning red: 'This is unbelievable! Every time I have to work early hours, they are having a bash!' **[VU – I]** He said, with his face turning red, that every time he has to work early hours, they are having a bash.
60	Carol, a taxi driver, was about to drive home when suddenly she saw a man by the road frantically waving his arms. Carol stopped and the man offered her double the fare if they make it to the airport within ten minutes. **[VF – D]** Being professional, she replied in a calm and firm voice: 'OK. I'll try my best, mate, but cannot work any miracles.' **[VF – I]** Being professional, she replied in a calm and firm voice that she will try her best, but that she cannot work any miracles. **[VU – D]** Being professional, she replied with a calm expression on her face: 'OK. I'll try my best, mate, but cannot work any miracles.' **[VU – I]** Being professional, she replied with a calm expression on her face, saying that she will try her best, but that she cannot work any miracles.	Carl, a taxi driver, was about to drive home when suddenly he saw a man by the road frantically waving his arms. Carl stopped and the man offered him double the fare if they make it to the airport within ten minutes. **[VF – D]** Being professional, Carl replied in a calm and firm voice: 'OK. I'll try my best, mate, but cannot work any miracles.' **[VF – I]** Being professional, Carl replied in a calm and firm voice that he will try his best, but that he cannot work any miracles. **[VU – D]** Being professional, Carl replied with a calm expression on his face: 'OK. I'll try my best, mate, but cannot work any miracles.' **[VU – I]** Being professional, Carl replied with a calm expression on his face, saying that he will try his best, but that he cannot work any miracles.

Appendix 3 Loudness stimuli

[D] = direct speech; [I] = indirect speech

The critical sentences are in bold and italic text (for acoustic analyses in oral reading).

Item	**Loud-speaking story**	**Quiet-speaking story**
1	Diane and Sharon entered the lobby during the intermission and purchased a glass of wine. The room was buzzing with hundreds of conversations about the performance. Diane spotted her sister's cheating ex-husband standing nearby with his new girlfriend and intended to publicly embarrass him. **[D]** She turned to Sharon and purposely increased her voice: 'Well! ***I can't believe he has the nerve to show his face in public after his awful behaviour***.' **[I]** She turned to Sharon and purposely increased her voice saying that ***she couldn't believe he had the nerve to show his face in public after his awful behaviour***.	The orchestra were playing a delicate, melancholy piece when Diane spotted her friend's cheating ex-husband sitting nearby. She was fuming to see him and nudged her sister, pointing him out to her subtly. **[D]** She turned and said in an undertone to Sharon: 'Well! ***I can't believe he has the nerve to show his face in public after his awful behaviour***.' **[I]** She turned and said in an undertone to Sharon that ***she couldn't believe he had the nerve to show his face in public after his awful behaviour***.
2	Katy's friends were having such a great time blasting cheesy Christmas music in her flat that she barely heard the knock at the door at 1am. She was expecting another friend and was surprised to find a policeman looking crossly over her shoulder at the party going on inside. **[D]** He looked very serious, demanding: 'Excuse me Madam. ***I suggest you keep the noise down as there's been a complaint from your neighbours***.' **[I]** He looked very serious, demanding that ***she should keep the noise down as there had been a complaint from her neighbours***.	Katy was reading in the library and had put ear plugs in to block out the other students studying nearby. She absentmindedly began to click her tongue, not realising that other people could hear her until a concerned librarian tapped her lightly on the shoulder. **[D]** He politely said: 'Excuse me Madam. ***I suggest you keep the noise down as there's been a complaint from your neighbours***.' **[I]** He politely said to her that ***she should keep the noise down as there had been a complaint from her neighbours***.
3	The construction site was bustling with activity and the shrill radio rang out in between the whirring of electric tools and reversing lorries. The builders and joiners had been working hard all morning so Bill, the foreman, gestured wildly to his team scattered around the site. **[D]** He yelled: 'Okay, ***I think it's time to have a cup of coffee***.' **[I]** He yelled that ***he thought it was time to have a cup of coffee***.	All that could be heard was the noise of picks on dry soil as the archaeologists painstakingly concentrated on the task. Bill was disappointed not to have found anything so far in spite of their conscientious work. He sighed softly and suggested taking a break to his co-worker. **[D]** He muttered: 'Okay, ***I think it's time to have a cup of coffee***.' **[I]** He muttered that ***he thought it was time to have a cup of coffee***.
4	On one terrifying night during WWII, the bombs were dropped over the city by the enemy and the troops could hear the explosions as they took cover in ruined buildings. Brian put his hands over his ears and screamed uncontrollably as a blast nearby made the walls of their hideout shake violently. **[D]** He cried: 'Bloody hell! ***It's really close! We should try not to move***.' **[I]** He cried that ***it was really close and that they should try not to move***.	It was a perfectly calm day and Sarah and Brian drifted along the coast in their sea kayak, spotting local birds and wildlife. As they rounded a cliff, they came across a sea eagle nearby and Brian turned around and put his finger to his lips. **[D]** He mouthed: 'Bloody hell! ***It's really close! We should try not to move***.' **[I]** He mouthed that ***it was really close and that they should try not to move***.

5	The nursery was filled with lots of young children, screaming for attention from the busy childminder. One boy in particular had been told off many times for behaving disgracefully all day.	The three siblings knew they had to flee the city that night if they were going to escape persecution and crept along the dark lane. They climbed cautiously into a truck and the eldest made sure her youngest brother was completely hidden behind the wooden crates.
	[D] With her face turning red, she stressed to him: 'Listen. ***You have to be really good and not make a sound for the next few hours***.' **[I]** With her face turning red, she stressed to him that ***he had to be really good and not make a sound for the next few hours***.	**[D]** With her face white with worry, she lowered her voice: 'Listen. ***You have to be really good and not make a sound for the next few hours***.' **[I]** With her face white with worry, she lowered her voice saying that ***he had to be really good and not make a sound for the next few hours***.
6	Harry dodged the flaming dragon, hearing it roaring over the game's soundtrack, and reached the next level of the computer game. Harry whooped euphorically over the sound of his friends shouting out congratulations.	The class were supposed to be revising independently but all the boys were passing around a new handheld computer game under the teacher's nose. Harry almost whooped when he managed to complete it and he leant over to show his friends secretly below the table.
	[D] Grinning from ear to ear, he boasted smugly: '***I told you that I'd manage it and I beat all of your scores***.' **[I]** Grinning from ear to ear, he boasted smugly that ***he had told them he would manage it and he had beaten all of their scores***.	**[D]** Winking, he said under his breath: '***I told you that I'd manage it and I beat all of your scores***.' **[I]** Winking, he said under his breath that ***he had told them that he would manage it and he had beaten all of their scores***.
7	John and Fiona had been dancing near the stage for hours and when the heavy metal concert ended, their ears were ringing. They made their way to the VIP area, hoping to catch a glimpse of the band, but were stopped by a stern steward guarding the area.	The carriage was virtually deserted at 11pm and John and Fiona felt themselves falling asleep with the motion of the train. They woke up groggily to see the inspector standing over them, gently prodding their shoulders.
	[D] Folding his arms across his chest, he boomed: '***I'm sorry but I need to see your tickets***.' **[I]** Folding his arms across his chest, he boomed that ***he was sorry but he needed to see their tickets***.	**[D]** He explained with an apologetic shrug: '***I'm sorry but I need to see your tickets***.' **[I]** He explained with an apologetic shrug that ***he was sorry but he needed to see their tickets***.
8	The nightclub was very busy when an irritated drunk man tried to punch a member of the bar staff. As the police were called, two burly bouncers approached him menacingly from both directions.	The movie premiere was going well with the audience sitting in tense anticipation as the thriller reached a disquieting stage. Two ushers appeared noiselessly in the aisle beside Sam and Fiona and motioned to a man illegally recording the new film.
	[D] Brusquely blocking his way, they said forcefully: 'Excuse me sir. ***We'd like you to accompany us upstairs to the office***.' **[I]** Brusquely blocking his way, they said forcefully that ***they would like him to accompany them upstairs to the office***.	**[D]** They leant towards the man and said in a hushed voice: 'Excuse me sir. ***We'd like you to accompany us upstairs to the office***.' **[I]** They leant towards the man and said in a hushed voice that ***they would like him to accompany them upstairs to the office***.
9	The raucous cheers rang out into the stadium, the fans exploding with joy as the ball was headed into the back of the net by the striker. Steven and his friends were hugging each other, stomping and cheering.	It had been a tense Wimbledon match ending in a tie and the game would be decided by sudden death. As they waited for the UK duo to serve, the spectators were warned not to distract the players.
	[D] Dancing in his seat, he was cheering joyfully: 'Come on. ***If we win this match, we might have a shot at the trophy***.' **[I]** Dancing in his seat, he was cheering joyfully that ***if they won the match, they might have a shot at the trophy***.	**[D]** Steven leaned in excitedly to his friend and murmured: 'Come on. ***If we win this match, we might have a shot at the trophy***.' **[I]** Steven leaned in excitedly to his friend and murmured that ***if they won the match, they might have a shot at the trophy***.

10	The bars had just closed at 1am and the roads were full of people merrily making their way home. As Helen was walking in the other direction, Tara suddenly remembered something and tried to catch her attention.	Helen and Tara entered their flat, making sure not to slam the door and wake up their flatmates. Helen removed her heels and tip-toed across the hall when Tara tapped her lightly on the shoulder.
	[D] She called to her: 'Oi! Don't forget! ***Paul's birthday party is on Friday so we'll need to go shopping***.' **[I]** She called to her that ***Paul's birthday party was on Friday and they would need to go shopping***.	**[D]** She discreetly said: 'Oi! Don't forget! We've got ***Paul's birthday party on Friday so we'll need to go shopping***.' **[I]** She discreetly said that they would have ***Paul's birthday party on Friday so they would need to go shopping***.
11	It was the first day back at school and all the children were very excited to see each other again. The journey home was chaos for the poor bus driver with pupils standing on their seats and shouting to friends.	As usual, the office was vacant and the two night watchmen glanced at the CCTV screens sleepily from time to time. Both men gulped as they felt a cold hard blade pressed to their throats and soft footsteps on the floor behind them.
	[D] Finally the driver had enough, pulled over and exclaimed angrily: 'Listen. ***I want you to stay in your seat and keep quiet if you want to get home tonight***.' **[I]** Finally the driver had enough, pulled over and exclaimed angrily that ***he wanted them to stay in their seats and keep quiet if they wanted to get home tonight***.	**[D]** A ghostly voice breathed: 'Listen. ***I want you to stay in your seat and keep quiet if you want to get home tonight***.' **[I]** A ghostly voice breathed that ***he wanted them to stay in their seats and keep quiet if they wanted to get home tonight***.
12	The salsa teacher threw them straight into dancing, barking out orders and dragging them to the front of the room. Susan and Jane were partnering each other and having a great time twirling to the fast-paced Latin beats.	Susan and Jane were growing restless as the meeting dragged on past five o'clock and kept shooting sneaky looks at each other. As their manager turned his back, Susan rolled her eyes and leaned over to Jane.
	[D] Giggling at her dizzy head, Susan shrieked: 'Oh dear. ***I think I'm going to need a drink after this***.' **[I]** Giggling at her dizzy head, Susan shrieked that ***she thought she was going to need a drink after this***.	**[D]** She said covertly: 'Oh dear. ***I think I'm going to need a drink after this***.' **[I]** She said covertly that ***she thought she was going to need a drink after this***.
13	All that Gary and Shona could hear was the whistling of the wind and the rain making hard thuds as it hit their jackets. They held onto their hoods as they struggled against the fierce gusts and tried to find the way back to their hotel in the town.	Gary and Shona took a late night stroll in a foreign city, carrying a large amount of money. When they took a wrong turn, Gary was concerned about being mugged and was conscious of the few people around them.
	[D] Gary couldn't stand it any longer and roared above the wind: '***I should go and ask for directions from someone in that restaurant***.' **[I]** Gary couldn't stand it any longer and roared above the wind that ***he should go and ask for directions from someone in that restaurant***.	**[D]** Gary moved closer so as not to be overheard and said in Shona's ear: '***I should go and ask for directions from someone in that restaurant***.' **[I]** Gary moved closer so as not to be overheard and said in Shona's ear that ***he should go and ask for directions from someone in that restaurant***.
14	The shop floor's speakers thudded with a new drum and bass track as Laura looked through the old records and found an early copy of her favourite album. She didn't want to risk leaving it in the shop when her mobile rang so she put her hand over her ear as she answered.	Laura returned to collect her developing slides and found the lab full of other scientists working independently. Her mobile began to vibrate, raising eyebrows from her co-workers and she grimaced apologetically.
	[D] She said loudly into the mouthpiece: '***I'm a little bit busy at the moment but I'll call you back in ten minutes***.' **[I]** She said loudly into the mouthpiece that ***she was a little bit busy at the moment but she would call them back in ten minutes***.	**[D]** She quietly answered: '***I'm a little bit busy at the moment but I'll call you back in ten minutes***.' **[I]** She quietly answered that ***she was a little bit busy at the moment but she would call them back in ten minutes***.

15	The thunderous drone of the boat's motor added to the splashing of the waves and the wind in their ears as Ian and David made their way out of the harbour. Steering the boat, Ian could just make out David saying something about sandwiches below deck.	Ian and David liked going to the cinema but resented paying for the overpriced refreshments so they had sneaked in their own. David was rustling through his bag as the film began to play and the room fell silent.
	[D] Ian tried to see into the hull and called down: '***If you're looking for the crisps and sandwiches, they're in the red rucksack***.' **[I]** Ian tried to see into the hull and called down that ***if he was looking for the crisps and sandwiches, they were in the red rucksack***.	**[D]** Ian leaned over and whispered: '***If you're looking for the crisps and sandwiches, they're in the red rucksack***.' **[I]** Ian leaned over and whispered that ***if he was looking for the crisps and sandwiches, they were in the red rucksack***.
16	Leanne's son had promised her that he would tidy his room when he got home from school. She was furious to find him watching a film, his speakers blaring and his room still a mess.	On Sunday evening, Leanne saw her son reading comic books in his room which was still messy despite his promise. As she spoke to him in his room, she was furious but tried not to disturb her youngest daughter sleeping next door.
	[D] She lunged to turn the TV off and exploded: 'That's it! ***You're going to be grounded until this room is sparkling***.' **[I]** She lunged to turn the TV off and exploded that ***he was going to be grounded until his room was sparkling***.	**[D]** She pulled his ear and hissed: 'That's it! ***You're going to be grounded until this room is sparkling***.' **[I]** She pulled his ear and hissed that ***he was going to be grounded until his room was sparkling***.
17	The gym was full of people training, running, kicking and punching to heavy bass pumping from the ceiling speakers. Helen had been trying to improve her kickboxing for weeks and her trainer was pushing her to the limit.	The teacher set the class to work silently on their individual projects and began to mark their essays. He was shocked at the unexpectedly low standard of Helen's work and asked her to speak to him outside the class to stop the other pupils hearing.
	[D] He yelled at her: 'Come on. ***I know you can do better than this so I want you to repeat the exercise***.' **[I]** He yelled at her that ***he knew that she could do better than that so he wanted her to repeat the exercise***.	**[D]** He said in a reassuring voice: 'Come on. ***I know you can do better than this so I want you to repeat the exercise***.' **[I]** He said in a reassuring voice that ***he knew that she could do better than that so he wanted her to repeat the exercise***.
18	In her bedroom, Annie put on some cheerful music while she was packing her suitcase for her holiday to Spain. Downstairs, Patrick looked at his watch and realised that they would have to leave soon to catch their flight.	The library was full of students studying seriously for their final exams. Patrick looked at his watch and realised that he and Annie should leave for their exam.
	[D] He stood at the bottom of the stairs and shouted: 'Come on, darling. ***We should leave in about half an hour if we don't want to be late***.' **[I]** He stood at the bottom of the stairs and shouted that ***they should leave in about half an hour if they didn't want to be late***.	**[D]** Trying not to disturb the other students nearby, he whispered to Annie: 'Come on, darling. ***We should leave in about half an hour if we don't want to be late***.' **[I]** Trying not to disturb the other students nearby, he whispered to Annie that ***they should leave in about half an hour if they didn't want to be late***.
19	Daniel was racing his new motor wildly in the street and he lost control of the vehicle, rushing into a lorry. In the hospital, he knew his father would be incredibly angry with him for speeding.	Daniel had been very lucky to survive the car accident when a drunk driver hit his car. His father looked heartbroken when he saw his son in the hospital bed for the first time.
	[D] When his father arrived, he slammed his fist into the hospital door and roared: 'Jesus Christ! ***You're damn lucky to be alive today***.' **[I]** When his father arrived, he slammed his fist into the hospital door and roared that ***he was damn lucky to be alive today***.	**[D]** His face was white as he said in a sad voice: 'Jesus Christ! ***You're damn lucky to be alive today***.' **[I]** His face was white as he said in a sad voice that ***he was damn lucky to be alive today***.

20	Alex was working in a busy antiques store on the local high street. At the most hectic point of the morning, an old man asked her repeatedly about one of their tables but wasn't listening carefully to her response.	The two art dealers were searching through the small, dusty antiques shop when one of them spotted a bargain. He didn't want to draw the owner's attention to the true rarity and expense of the piece.
	[D] She couldn't make herself any clearer as she exclaimed: 'Listen. ***The table is £2500 and is from the late 18th century***.' **[I]** She couldn't make herself any clearer as she exclaimed that ***the table was £2500 and was from the late 18th century***.	**[D]** He turned to his brother and gestured, discreetly saying: 'Listen. ***The table is £2500 and is from the late 18th century***.' **[I]** He turned to his brother and gestured, discreetly saying that ***the table was £2500 and was from the late 18th century***.
21	The Masterchef studio kitchen was filled with the clanging of pots and pans and the sizzling of food. The loudspeaker was activated in the hectic kitchen for the presenter to make an announcement to the three finalists.	Heather had missed her class exam and had to sit the test alone under the supervision of her English teacher. As the final half an hour approached, she was focused entirely on her work and the teacher kindly reminded her of the time.
	[D] He tapped the microphone and declared with authority: 'OK. ***It's quarter past one so there's half an hour left until the interval***.' **[I]** He tapped the microphone and declared with authority that ***it was quarter past one so there was half an hour left until the interval***.	**[D]** He cleared his throat, gently saying: 'OK. ***It's quarter past one so there's half an hour left until the interval***.' **[I]** He cleared his throat, gently saying that ***it was quarter past one so there was half an hour left until the interval***.
22	James was annoyed when the documents he had asked to be sent were returned to the company in the post. He was in a foul mood and stormed through to his assistant's office, not bothering to knock on the door.	The woods were still and motionless as the car pulled into the lane and turned off its engine. A knock on the glass made James start and he rolled down his window but couldn't see the agent's face.
	[D] Banging the package on the desk, James angrily erupted: '***Make sure this gets to the right person***.' **[I]** Banging the package on the desk, James angrily erupted that he should ***make sure that it got to the right person***.	**[D]** Handing him the package, James muttered: '***Make sure this gets to the right person***.' **[I]** Handing him the package, James muttered that he should ***make sure that it got to the right person***.
23	Erica had taken her daughter to an indoor adventure play area where there was an assault course, a ball pool and inflatable toys. After an hour of watching her play with the other children, she motioned energetically in the crowd to get her attention.	Erica loved going for an aromatherapy massage and felt her muscles soothe as the hour went on. She felt herself drifting off with the smell of the oils and the relaxing music when the masseuse turned down the CD.
	[D] Erica shouted out: '***I'm afraid that's time up. I'll meet you at reception***.' **[I]** Erica shouted out that ***she was afraid that it was time up and that she would meet her at reception***.	**[D]** She placed her hand lightly on her shoulder, saying soothingly: '***I'm afraid that's time up. I'll meet you at reception***.' **[I]** She placed her hand lightly on her shoulder, saying soothingly that ***she was afraid that it was time up and that she would meet her at reception***.
24	Everyone in Kelly's living room was intrigued as the documentary showed the plastic surgery taking place. However, as the camera zoomed in on the gory scene, the living room exploded with utter disgust.	An ambulance siren could be heard in the distance as Kelly edged past distressed neighbours to see the accident. She saw the man's body lying in the road, surrounded by the eerily motionless crowd.
	[D] Kelly couldn't believe what she was seeing and yelled: 'Oh my god. ***I've never seen anything so disturbing***.' **[I]** Kelly couldn't believe what she was seeing and yelled that ***she had never seen anything so disturbing***.	**[D]** Kelly covered her mouth in shock and murmured: 'Oh my god. ***I've never seen anything so disturbing***'. **[I]** Kelly covered her mouth in shock and murmured that ***she had never seen anything so disturbing***.

Appendix 4 speaking rate stimuli

[D] = direct speech; [I] = indirect speech; [Q] = comprehension question (for Experiment 10 only)

The bold and italic texts in direct and indirect speech quotations are the critical regions.

Item	Fast-speaking story	Slow-speaking story
1	It was 11am in the morning when the fire alarm went off. Hearing people running down the corridors, Mary grabbed her jacket and burst into Peter's office next door.	It was 11am in the morning when the fire alarm went off. Knowing that this was just a test, Mary put on her jacket and walked into Peter's office next door.
	[D] She shouted: 'Peter, quick, we have ***to leave immediately because the building is on fire***!' **[I]** She urged Peter ***to leave immediately because the building was on fire***.	**[D]** Tongue-in-cheek, she said: 'Peter, quick, we have ***to leave immediately because the building is on fire***!' **[I]** Tongue-in-cheek, she told Peter ***to leave immediately because the building was on fire***.
	Confused, Peter replied: 'Wasn't there an e-mail about a fire alarm test this morning?'	Peter just replied: 'Very funny – I've seen the e-mail about the fire alarm test as well!'
	[Q] Did Mary think it was a real fire?	**[Q]** Did Mary think it was a real fire?
2	It was a Sunday morning during the summer vacation. At the airport, Cheryl and Gareth were boarding the flight to Paris for their honeymoon and there were only 5 minutes left before the plane was taking off. Cheryl started panicking because she couldn't find her passport.	It was a Sunday morning during the summer vacation. Cheryl and Gareth finally found the time to sort some personal documents and other clutter that had been piling up on the kitchen table. Cheryl was looking for her passport and became confused.
	[D] Hastily, she said to Gareth: 'Oh no! ***I am pretty sure it was in that sheath, but I cannot find it anymore***!' **[I]** Hastily, she said to Gareth that ***she was pretty sure it was in that sheath, but she couldn't find it anymore***.	**[D]** She turned to Gareth and said: 'Hmm, that's weird. ***I am pretty sure it was in that sheath, but I cannot find it anymore***!' **[I]** She turned to Gareth and said that it was weird because ***she was pretty sure it was in that sheath, but she couldn't find it anymore***.
	Gareth went straight back to the duty-free shop, and luckily, he found the passport by the till.	Gareth looked under a pile of bills, and luckily, he found her passport there.
	[Q] Did they manage to catch the flight?	**[Q]** Did Cheryl need her passport at the time?
3	It was a typical British day, rainy and gloomy. Sixteen year-old pianist Bobby was going to play in the quarter-final of a local talent competition. He was extremely nervous before his performance.	It was a typical British day, rainy and gloomy. At Glasgow Royal Infirmary, an old man was dying, and too weak to sit up. His family members were sitting around the bed, feeling sad. He wanted to say something, so his daughter placed a cushion under his head.
	[D] His mother encouraged him but he was all shaking and said: 'No! I can't do it! ***This is the end of the journey because it is unlikely that I will make it this time***.' **[I]** His mother encouraged him but he was all shaking and said that he couldn't do it and that ***it was the end of the journey because it was unlikely that he would make it this time***.	**[D]** Slowly, he looked around and said: 'I'm grateful you're all here. ***This is the end of the journey because it is unlikely that I will make it this time***.' **[I]** Slowly, he looked around and said that he was grateful for their coming and that ***it was the end of the journey because it was unlikely that he would make it this time***.
	His mother tried to calm him down, saying that it's not the winning that counts, but the taking part.	Then he closed his eyes and everyone burst into tears.
	[Q] Was Connie anxious about her coming performance in the quarter-final?	**[Q]** Was the man too weak to open his mouth?

4	At the beach, William and Phillip were enjoying the sunshine and a fresh sea breeze. Suddenly, William noticed a piece of wood floating on the water with something lying on it that looked like an unconscious person. **[D]** Without hesitation, he exclaimed: 'Phil, quick! We need ***to call for help because there might be someone in trouble***!' **[I]** Without hesitation, he exclaimed that they needed ***to call for help because there might be someone in trouble***. Phillip immediately grasped his mobile phone and dialled 999. **[Q]** Was there a person in danger?	At the beach, William and Phillip were enjoying the sunshine and a fresh sea breeze. They were discussing the hypothetical case of what to do if they saw someone drowning. **[D]** William thought there would be little one could do, but Phil objected: 'Look. The most important thing is ***to call for help because there might be someone in trouble***.' **[I]** William thought there would be little one could do, but Phil objected that the most important thing would be ***to call for help because there might be someone in trouble***. After a while, the conversation moved on to a different topic. **[Q]** Was there a person in danger?
5	It was at that peaceful house in Dumbarton where Mrs. Harris was preparing lunch for her children. Suddenly, the phone rang. It was Mr. Harris's business partner, urgently informing Mrs. Harris that her husband has had a car crash this morning. **[D]** He spoke hastily: 'Mrs. Harris, please come immediately! Your husband ***was involved in a collision with a heavy goods vehicle and is now in hospital***!' **[I]** He hastily told Mrs. Harris to come immediately because Mr. Harris ***was involved in a collision with a heavy goods vehicle and was now in hospital***. Mrs. Harris was shocked, and headed for the hospital right away. **[Q]** Was Mr. Harris dead?	It was at that peaceful house in Bristol where PCs Bryce and Rawlinson had some sad news to impart. The owner of the house, Mr. Harris, was killed in a car crash earlier this morning. Paramedics had rushed on the scene in time, but they couldn't save him. **[D]** With sorrow, Bryce told Mrs. Harris: 'Mrs. Harris, I am afraid that your husband ***was involved in a collision with a heavy goods vehicle and he is now in hospital***.' **[I]** With sorrow, Bryce told Mrs. Harris that her husband ***was involved in a collision with a heavy goods vehicle and was now in hospital***. Mrs. Harris sobbed. Apparently, she could not believe it. **[Q]** Was Mr. Harris dead?
6	On Valentine's Day, Simon at long last worked up the courage to propose to Emma. Yet to his horror, she refused. He felt humiliated and blamed Emma for her brusque reply. **[D]** Emma fought back: 'Come on, Simon! How could you blame me for this? ***I am simply not ready for such a commitment***!' **[I]** Emma fought back, saying that Simon could not really blame her for that and that ***she was simply not ready for such a commitment***. Simon was speechless, realising that he was humiliating himself. **[Q]** Was Simon's proposal successful?	On Valentine's Day, Simon at long last worked up the courage to propose to Emma. Yet to his horror, she refused. **[D]** He asked for the reason and Emma replied calmly: 'You are a really nice bloke, Simon, but ***I am simply not ready for such a commitment***.' **[I]** He asked for the reason and Emma replied calmly that although she thought he was a really nice bloke, ***she was simply not ready for such a commitment***. Simon was upset for a moment but determined not to give up. **[Q]** Was Simon's proposal successful?

7	Alice was looking for pheasant paintings for her art history dissertation. Today her boyfriend John bumped into a weird-looking man with greasy hair and thick glasses who, amazingly, claimed to possess a sizeable collection of what Alice was looking for.	Alice was selling paintings in a small shop downtown. One day, a weird-looking man with greasy hair and thick glasses came into her shop.
	[D] John was excited and phoned Alice immediately: 'Alice, you wouldn't believe it!.. I just met ***one of the world's keenest collectors of pheasant paintings***!' **[I]** John was excited and phoned Alice immediately, saying that he just met ***one of the world's keenest collectors of pheasant paintings***.	**[D]** He looked around and said nonchalantly: 'You may be surprised to learn that I am ***one of the world's keenest collectors of pheasant paintings***.' **[I]** He looked around and said nonchalantly that she might be surprised to learn that he was ***one of the world's keenest collectors of pheasant paintings***.
	'Really?!' Alice couldn't believe her ears.	Alice replied: 'Uhm.. I don't think we have any at the moment.'
	[Q] Was the man interested in pheasant paintings?	**[Q]** Was the man interested in pheasant paintings?
8	On the train to London, Olivia was very nervous about her upcoming job interview. Looking through the window, she reached for a cigarette without thinking, just to calm herself down.	Olivia and her 5 year old daughter were on a train to London. Looking through the window, a thoughtless youth lit a cigarette and was immediately reprimanded by other passengers. Olivia's daughter was curious and asked why one cannot smoke on the train.
	[D] Instantly, a middle-aged man next to her complained: 'Oi! Stop it! ***Smoking on public transport is illegal***!' **[I]** Instantly, a middle-aged man next to her complained that she should stop that because ***smoking on public transport was illegal***.	**[D]** Olivia smiled and answered: 'Well, since it may harm other passengers' health, ***smoking on public transport is illegal***.' **[I]** Olivia smiled and answered that since it may harm other passengers' health, ***smoking on public transport is illegal***.
	Olivia put the cigarette back into the box and felt even more stressed after this incident.	The girl thought for a second and said: 'Why can't they just hold their breath?'
	[Q] Did Olivia smoke on the train?	**[Q]** Did Olivia answer her daughter patiently?
9	PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. Ella asked for an extension but Jim looked rather impatient.	PhD student Ella was summoned to her supervisor Jim's office to give a report on her current progress. Ella asked for an extension but Jim seemed to have a better idea.
	[D] Hastily, he said: 'No way! We really need ***to have those data in by next month for that conference***!' **[I]** Hastily, he said that they really needed ***to have those data in by next month for that conference***.	**[D]** He said: 'Actually…, it might be good ***to have those data in by next month for that conference***.' **[I]** He said that it might be good ***to have those data in by next month for that conference***.
	Ella was gutted because it meant that she would have to work like mad over the following weeks.	Ella understood his point and decided to work even harder over the following weeks.
	[Q] Was Ella's request for an extension rejected immediately?	**[Q]** Was Ella's request for an extension rejected immediately?

10	In the board room, assistant manager Craig was giving a presentation on how to improve the company's sales figures. A fellow manager was far from impressed and openly criticised Craig's public speaking skills.	In the board room, assistant manager Craig was giving a presentation on how to improve the company's sales figures. When he noticed that his colleagues looked far from impressed, he felt embarrassed.
	[D] Craig kind of lost his temper and was quick to defend himself: 'What do you expect?! ***I never had to give this kind of presentation before, so I am not very experienced***!'	**[D]** Craig hesitantly explained: 'Sorry, ***I never had to give this kind of presentation before, so I am not very experienced***.'
	[I] Craig kind of lost his temper and was quick to defend himself, saying that ***he never had to give this kind of presentation before, so he was not very experienced***.	**[I]** Craig hesitantly explained that ***he never had to give this kind of presentation before, so he was not very experienced***.
	Everyone else in the room went quiet.	Everyone else in the room went quiet.
	[Q] Was Craig apologetic about his performance?	**[Q]** Was Craig apologetic about his performance?
11	It was 5.30 pm and people were ready to leave the office. Elaine seemed to be rather agitated, cursing a lot in front of her screen. Steven went to Elaine's desk asked and what's wrong.	It was 5.30 pm and people were ready to leave the office. Steven went to Elaine's desk and saw her slumped in her chair.
	[D] Elaine complained: 'Just look at this! ***The amount of admin is killing me at the moment, and there seems to be no end to it in sight***!'	**[D]** He asked what's wrong, and Elaine replied with a listless expression on her face: 'I'm totally worn out. ***The amount of admin is killing me at the moment, and there seems to be no end to it in sight***.'
	[I] Elaine complained that ***the amount of admin was killing her at the moment, and that there seemed to be no end to it in sight***.	**[I]** He asked what's wrong, and Elaine replied with a listless expression on her face that ***the amount of admin was killing her at the moment, and that there seemed to be no end to it in sight***.
	Steven knew exactly what she was talking about.	Steven knew exactly what she was talking about.
	[Q] Was Elaine agitated about admin?	**[Q]** Was Elaine agitated about admin?
12	It was half past seven and people were unusually busy at the Aberdeen restaurant. The chef could hardly cope with the pressure of serving a bus-load of tourists who had just arrived. After a couple of complaints, the restaurant manager went berserk.	It was half past seven and people were unusually busy at the Aberdeen Cooking College. The students had just finished their final exam and were awaiting the results. In front of Mike's dishes, the examiners frowned.
	[D] He dashed straight into the kitchen and shouted: 'Is this how to treat our customers!? ***The fish is burnt, the vegetables overcooked, and the soup is actually stone cold***!'	**[D]** After exchanging their views, the head examiner finally announced a verdict: 'Mike, we really appreciate the efforts you've made, but ***the fish is burnt, the vegetables overcooked, and the soup is actually stone cold***.'
	[I] He dashed straight into the kitchen and shouted whether this was how to treat their customers, pointing out that ***the fish was burnt, the vegetables overcooked, and the soup was actually stone cold***.	**[I]** After exchanging their views, the head examiner finally announced a verdict, saying that although they really appreciated the efforts Mike had made, ***the fish was burnt, the vegetables overcooked, and the soup was actually stone cold***.
	Everything went instantly quiet in the kitchen, except for the occasional bubbling noises from the cooker.	Mike felt upset because this could only mean that he had failed.
	[Q] Was the manager angry with the chef?	**[Q]** Did the examiners fail Mike immediately?

13	Jenny's fiancé Ben had a bad motorcycle accident which was likely to leave him paraplegic for the rest of his life. Calling from the hospital, he said farewell to Jenny, anticipating that she might leave him because of his condition.	Jenny and Ben were having a candle light dinner to celebrate the first anniversary of their relationship. Yet at some point, Jenny hesitantly mentioned that because of her new job she had to move to another city.
	[D] Jenny was crying and shouting: 'No! ***I'm not leaving you! I want to be with you whatever happens because I love you***!'	**[D]** After a sip of champagne, Ben held her hands and said very calmly: 'Don't worry, darling. ***I'm not leaving you. I want to be with you whatever happens because I love you***.'
	[I] Jenny was crying and shouting that ***she was not leaving him and that she wanted to be with him whatever happened because she loved him***.	**[I]** After a sip of champagne, Ben held her hands and said very calmly that ***he was not leaving her and that he wanted to be with her whatever happens because he loved her***.
	Jenny's reassurance came as a great relief.	Ben's reassurance came as a great relief.
	[Q] Did Ben have a motorcycle accident?	**[Q]** Did Jenny have a new job?
14	Shauna was a young, talented actress and singer. She was desperate for a part in the new musical project but her boyfriend was ill and blamed her for not looking after him.	Theatrical agent Peter was visited by one of his clients. Shauna was a young, talented actress and singer. She was desperate for a part in the new musical project.
	[D] Shauna was angry with his attitude, blustering: 'Don't be ridiculous! ***I am doing this because without this job, I can't even pay for the electricity bill***!'	**[D]** Shauna hummed and hawed: 'It is embarrassing to say, but…, ***I am doing this because without this job, I can't even pay for the electricity bill***!'
	[I] Shauna was angry with his attitude, blustering that ***she was doing this because without this job, she couldn't even pay for the electricity bill***.	**[I]** Shauna hummed and hawed with embarrassment, saying that ***she was doing this because without this job, she couldn't even pay for the electricity bill***.
	Her boyfriend was speechless and realised how childish he was.	Peter was moved and promised he would try his best to help her.
	[Q] Did Shauna think her boyfriend was sensible?	**[Q]** Did Shauna need money?
15	Harry and Leona were having a traditional Sunday roast. Harry was constantly talking with his mouth full. Suddenly, he stopped and started choking, with his face turning bright red.	Red Cross nurse Leona was teaching the public how to make an emergency call. She first introduced a list of important information that should be provided.
	[D] Leona was panicking and dialled 999: 'Help! ***My husband is choking on his food and I actually don't know what to do! It looks rather serious***!'	**[D]** Then she demonstrated how NOT to do it by mocking a nonchalant woman saying: 'It seems as if ***my husband is choking on his food and I actually don't know what to do. It looks rather serious***…'
	[I] Leona was panicking and dialled 999, saying that ***her husband was choking on his food and that she actually didn't know what to do, adding that it looked rather serious***.	**[I]** Then she demonstrated how NOT to do it by mocking a nonchalant woman saying that it seemed as if ***her husband was choking on his food and that she actually didn't know what to do, adding that it looked rather serious***.
	The person at the other end instantly gave her some emergency instructions.	Some people in the audience started to chuckle.
	[Q] Was Harry having a soup?	**[Q]** Was Leona a school teacher?

16	It was the last day of John Lewis's summer sale when two girls were browsing around. Helen was very fond of a black skirt and shrieked when she found it was heavily discounted.	It was the last day of John Lewis's summer sale when two girls were browsing around. Helen felt almost heartbroken when she found that her favourite skirt was heavily discounted – she and her friend Abbey were completely skint at the moment.
	[D] She could not wait to show it to her friend and shouted: 'Hey Abbey! Look at this one! Gorgeous isn't it!? ***It is exactly what I always wanted, and now it is even more than 50% off***!' **[I]** She could not wait to show it to her friend Abbey. Excited, she asked her to have a look, shouting that ***it was exactly what she always wanted, and now it was even more than 50% off***.	**[D]** With a deep-drawn sigh, she said: 'Look, Abbey. Gorgeous isn't it… ***It is exactly what I always wanted, and now it is even more than 50% off***…' **[I]** With a deep-drawn sigh, she said to Abbey that ***it was exactly what she always wanted, and now it was even more than 50% off***.
	Abbey agreed: 'I'm sure it will look great on you.'	Abbey tried to comfort her: 'I'm sure there will be other good offers next time round.'
	[Q] Was Helen excited about the skirt?	**[Q]** Was Helen fond of the skirt?
17	It was 1.55pm when the Riley's were about to head off to the train station to collect granddad.	It was 1.55pm when the Riley's were about to head off to the train station to collect granddad. Fourteen year-old Karen was asking why they went so early, given that the train was due to arrive at 3pm.
	[D] Mrs Riley was panicking because she couldn't find the car keys: 'Jesus! We've got only five minutes to go! ***If we don't arrive in time, granddad might easily get lost, considering how forgetful he sometimes is***!' **[I]** Mrs Riley was panicking because she couldn't find the car keys. She realized they've got only five minutes to go, and said that ***if they didn't arrive in time, granddad might easily get lost, considering how forgetful he sometimes was***.	**[D]** Her mum replied with a smile: 'Well, ***if we don't arrive in time, granddad might easily get lost, considering how forgetful he sometimes is***.' **[I]** Her mum replied with a smile, saying that ***if they didn't arrive in time, granddad might easily get lost, considering how forgetful he sometimes was***.
	'Don't panic,' her husband said, 'I've already called a cab.'	Karen remained defiant: 'But it is Sunday, and there is almost no traffic!'
	[Q] Will the train arrive at 2pm?	**[Q]** Will the train arrive at 2pm?
18	At Glasgow's Royal Infirmary, surgeons and nurses were trying their best to save a man's life. The minutes kept ticking away and the head surgeon looked concerned.	At Glasgow's Royal Infirmary, the students were shown some important medical emergency procedures.
	[D] He was shouting: 'Quick! ***Set up the life-support machine and stop the bleeding***!' **[I]** He ordered to quickly ***set up the life-support machine and stop the bleeding***.	**[D]** The tutor gave precise instructions: 'When the patient comes in, first ***set up the life-support machine and stop the bleeding***!' **[I]** The tutor gave precise instructions, saying that when the patient comes in, first they should ***set up the life-support machine and stop the bleeding***.
	Nurses followed his instructions promptly but the man showed little sign of life.	The students were taking notes and listened carefully.
	[Q] Was it important to set up the life-support machine?	**[Q]** Was it important to set up the life-support machine?

19	It was Tuesday, 6th June 1944. At Omaha Beach, allied soldiers braved artillery fire, mortar attacks and all-round danger. Enemy gunfire poured at the approaching allies and kept them at mid-distance.	At Omaha Beach, allied soldiers braved artillery fire, mortar attacks and all-round danger. In the days before the invasion, the assault had been meticulously planned at allied headquarters.
	[D] The captain of the 1st Division spurred his soldiers on: 'There is not much time! ***We must silence those machine guns before the reinforcements arrive***!' **[I]** The captain of the 1st Division spurred his soldiers on, shouting that there was not much time, and that ***they must silence those machine guns before the reinforcements arrive***.	**[D]** Back then, the officer in charge explained: '***We must silence those machine guns before the reinforcements arrive***.' **[I]** Back then, the officer in charge explained that ***they must silence those machine guns before the reinforcements arrive***.
	Then he jumped up and charged forward.	Obviously, the success of the whole operation depended on that.
	[Q] Should the machine guns be silenced?	**[Q]** Should the machine guns be silenced?
20	A heavily armed robbery took place earlier this morning. Within five minutes, five police cars showed up at the scene, but the busy streets made it difficult to locate the getaway culprits.	The West End festival was coming soon and the organisers were discussing details of the live television coverage. John considered setting up several cameras, but the busy streets made it difficult to locate the parade.
	[D] In the heat of the situation, the chief made a prompt decision and ordered: 'Block the entire West End! ***We need a helicopter to cover the whole area***!' **[I]** In the heat of the situation, the chief made a prompt decision and ordered to block the entire West End, adding that ***they needed a helicopter to cover the whole area***.	**[D]** Pointing at the map, John said: 'Hmm, not sure whether street cameras would work. Perhaps ***we need a helicopter to cover the whole area***.' **[I]** Pointing at the map, John said that he was not sure whether street cameras would work, adding that perhaps ***they needed a helicopter to cover the whole area***.
	Then he jumped into his car and headed for the nearest exit.	Unfortunately, this plan was well out of budget.
	[Q] Was a helicopter needed?	**[Q]** Was a helicopter needed?
21	Throughout human history, wars have caused many deaths. Today, United Nations Peacekeeping Forces were ambushed by Taliban forces in northern Afghanistan. They returned fire immediately, but suffered heavy losses.	Throughout human history, wars have caused many deaths. Recently, Tom was interviewing a 75 year-old Vietnam veteran. When Tom asked about the situation in the Battle of Hue, the old soldier paused for a while to recall his memories.
	[D] The officer in charge hastily called for reinforcements, shouting: '***We are under attack, facing mortar fire and rocket propelled grenades***!' **[I]** The officer in charge hastily called for reinforcements, shouting that ***they were under attack, facing mortar fire and rocket propelled grenades***.	**[D]** Then he slowly replied: 'On the night of January 30, ***we were under attack, facing mortar fire and rocket propelled grenades***..' **[I]** Then he slowly replied that on the night of January 30, ***they were under attack, facing mortar fire and rocket propelled grenades***.
	As he finished his message, a shell exploded beside him.	His facial expression gave a clue of the terror he felt at the time.
	[Q] Was the story about the Vietnam war?	**[Q]** Was the story about the Afghanistan war?

22	It was almost midnight and moonlight dripped through the leaves on the pavements. In the workshop, the cruel overseer was driving his workers on, as he was under pressure to meet his targets.	It was almost midnight and moonlight dripped through the leaves on the pavements. Robert and Chloe just came back from a party. Robert paid the cab driver and wanted to go to sleep straight away. But Chloe was far from tired and fancied another drink.
	[D] When one of the workers started yawning, he instantly shouted: 'Oi, lazybones! ***The day has actually only just begun***!' **[I]** When one of the workers started yawning, he instantly shouted at her, calling her lazybones, and that ***the day had actually only just begun***.	**[D]** She patted Robert's cheek and said: 'Stop yawning, my friend. ***The day has actually only just begun***.' **[I]** She patted Robert's cheek and told him to stop yawning because ***the day had actually only just begun***.
	Although his methods may have been effective, he certainly was not very popular.	Robert rolled his eyes, but agreed to have one last pint somewhere.
	[Q] Was the overseer under pressure?	**[Q]** Did Chloe fancy another drink?
23	One morning, six-year-old Sara was taking a lot of time to get ready for school. Her dad got increasingly angry with her because the school bus was due to arrive soon.	One morning, six-year-old Sara was taking a lot of time to get ready. She was secretly hoping that she would miss the school bus if she kept dragging on.
	[D] He said: 'Sara! ***If you are not ready within five minutes, I'll have to drive you again***!' **[I]** He said to Sara that ***if she was not ready within five minutes, he'd have to drive her again***.	**[D]** Her dad could easily see what she was up to and said in a sarcastic tone: 'Just take your time.. ***If you are not ready within five minutes, I'm happy to drive you again***.' **[I]** Her dad could easily see what she was up to and said in a sarcastic tone that she could just take her time, adding that ***if she was not ready within five minutes, he'd be happy to drive her again***.
	Reluctantly, Sara complied and put on her coat.	Reluctantly, Sara complied and put on her coat.
	[Q] Was Sara taking a lot of time?	**[Q]** Was Sara taking a lot of time?
24	Millionaire Joseph was addicted to betting on horses. Tipped by a so-called 'insider', he recently placed an enormous bet, but shockingly, the horse had lost.	Millionaire Joseph was addicted to betting on horses. Tipped by a so-called 'insider', he recently placed an enormous bet, but shockingly, the horse had lost.
	[D] Angry with his informant, Joseph shouted furiously on the phone: 'Where did your bloody information come from!? ***That was a huge amount of money – almost one million pounds***!' **[I]** Angry with his informant, Joseph shouted furiously on the phone, asking where the information had come from, because ***that was a huge amount of money – almost one million pounds***.	**[D]** Realising that he was nearly bankrupt, Joseph slumped back into his chair and started talking to himself: '***That was a huge amount of money – almost one million pounds***…' **[I]** Realising that he was nearly bankrupt, Joseph slumped back into his chair and started talking to himself, repeatedly saying that ***it was a huge amount of money – almost one million pounds***.
	Then he realised that it was his wife speaking at the other end.	Suicidal thoughts were creeping up his mind.
	[Q] Did Joseph bet a small amount of money?	**[Q]** Did Joseph bet a small amount of money?

Appendix 5 Fixation Density Disparity

Fixation Density Disparity (FDD) is designed to quantify the fluency of the initial reading of a relatively long critical region (e.g., a sentence) in reading experiments using eye-tracking. The measure primarily captures the *spatial distribution* of fixations in a given region. As such, it is independent of the reading-time measures of interest – thus, it is unlikely to introduce a bias when used as an outlier criterion for reading-time analyses.

During initial reading of a region, **FDD** = $\begin{cases} 0 & (\text{if } f = 0), \\ \frac{F}{f} & (\text{if } f > 0) \end{cases}$

F = the number of *surrounding* fixations

where

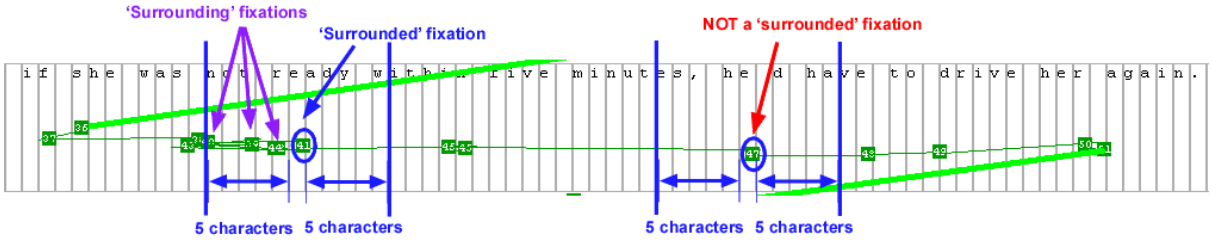
f = the number of *surrounded* fixations

A given fixation is characterised as a s*urrounded* fixation when there are other fixations within the vicinity of 5 characters to the left or right of that fixation (blue circled fixation in Example 1 below). These neighbouring fixations are *surrounding* fixations (fixations indicated by purple arrows in Example 1). The window of 5 characters to the left and right of a given fixation is motivated by a commonly reported average saccade size of 5-7 characters between fixations during normal reading (e.g., Morrison & Rayner, 1981)

A high FDD value indicates that there are clusters of spatially densely distributed fixations during initial reading of the critical region (e.g., due to frequent re-inspections of words), suggesting faltering reading (Example 1). By contrast, a low FDD value occurs when fixations are evenly distributed across the critical region, indicating reasonably fluent reading; an FDD of 0 suggests perfectly fluent reading (Example 2).

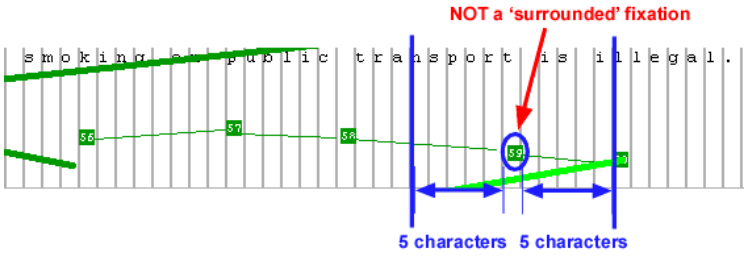
Note. *A cut-off FDD value of 2 (see outlier definition in paper) implies that the average go-past fixation per region may be surrounded by no more than two fixations within its 5-character vicinity.*

Example 1: Faltering reading (each green square represents a fixation; the lines between them represent saccades; each grey column represents a character position)



Fixation ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	total	FDD
***Number of surrounding* fixations at each given fixation**	1	1	3	6	5	3	4	3	4	1	1	0	1	1	1	1	F =36	2.40
***Surrounded* by other fixations? (0=no, 1=yes)**	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	f = 15	

Example 2: Fluent reading.



Fixation ID	1	2	3	4	5	total	FDD
***Number of surrounding* fixations at each given fixation**	0	0	0	0	0	F = 0	0
***Surrounded* by other fixations? (0=no, 1=yes)**	0	0	0	0	0	f = 0	

Reference

Morrison, R. E., & Rayner, K. (1981). Saccade size in reading depends upon character spaces and not visual angle. *Perception and Psychophysics, 30(4)*, 395-396.

Appendix 6 Supplementary measures (Experiment 10)

The *go-past time* measures reported in the paper are conceptually equivalent to the product of (i) the **numbers** of *go-past* fixations launched on the critical regions and (ii) the average **durations** of those *go-past* fixations. To examine whether the reported reading rate modulations in Experiment 10 were due to either (i) or (ii) (or indeed a combination of both), we conducted further analyses on these sub-measures.

(i) Numbers of fixations

Tables I and II below show the average **numbers** of *go-past* (respectively *no-regressions go-past*) fixations launched in each condition of Experiment 10. There was a significant *Context × Quoting Style* interaction in the number of *go-past* fixations launched ($F_1(1, 47) = 18.44$, $p < .001$; $F_2(1, 23) = 6.24$, $p = .02$) and, at least marginally, in the number of *no-regressions go-past* fixations launched ($F_1(1, 47) = 10.15$, $p < .002$; $F_2(1, 23) = 4.01$, $p < .06$).

Table I. Numbers of go-past fixations launched in Experiment 10 (with standard deviations in brackets). Also shown are results from 2-tailed paired sample *t*-tests examining the simple effects of *Context* and *Quoting Style*.

			Direct Speech	Indirect Speech	Quoting Style Contrast			
					By Subjects		By Items	
					t(47)	*p*	*t*(23)	*p*
Fast-speaking Context			9.5 (3.6)	10.3 (4.1)	**-2.72**	**.01**	**-2.54**	**.02**
Slow-speaking Context			10.1 (3.7)	9.7 (4.0)	0.72	.48	0.47	.64
Context Contrast	**By Subjects**	***t*(47)**	**-2.86**	**2.28**	Note: Significant contrasts are highlighted in bold			
		P	**.006**	**.03**				
	By Items	***t*(23)**	-1.78	1.57				
		P	.09	.13				

Table II. Numbers of no-regressions go-past fixations launched in Experiment 10 (with standard deviations in brackets). Also shown are results from 2-tailed paired sample *t*-tests examining the simple effects of *Context* and *Quoting Style*.

			Direct Speech	Indirect Speech	Quoting Style Contrast			
					By Subjects		By Items	
					t(47)	*p*	*t*(23)	*p*
Fast-speaking Context			9.1 (3.3)	9.7 (3.9)	**-2.84**	**.007**	**-2.14**	**.04**
Slow-speaking Context			9.7 (3.7)	9.3 (3.7)	0.98	.33	0.12	.91
Context Contrast	**By Subjects**	***t*(47)**	**-2.41**	**2.57**	Note: Significant contrasts are highlighted in bold			
		P	**.02**	**.01**				
	By Items	***t*(23)**	-1.62	1.21				
		P	.12	.24				

As we did in the main analyses in the paper, we ran the same analyses again, but with items 1, 23, and 24 removed. The *Context × Quoting Style* interaction remained reliable in numbers of *go-past* fixations ($F_1(1, 47) = 7.61$, $p < .01$, $F_2(1, 20) = 4.75$, $p = .04$) and, at least marginally, in numbers of *no-regressions go-past* fixations ($F_1(1, 47) = 4.82$, $p = .03$, $F_2(1, 20) = 3.25$, $p < .09$) – importantly, just as with the reading time measures in the paper,

the *Context* simple effect in the *indirect speech* condition (cf. Tables I and II) disappeared when those three items were removed (Tables III and IV).

Table III. Numbers of go-past fixations launched in Experiment 10 (with standard deviations in brackets) **after removing items 1, 23, and 24 from analysis**. Also shown are results from 2-tailed paired sample *t*-tests examining the simple effects of *Context* and *Quoting Style*.

			Direct Speech	Indirect Speech	Quoting Style Contrast			
					By Subjects		By Items	
					t(47)	*p*	*t*(20)	*p*
Fast-speaking Context			9.8 (3.7)	10.3 (4.2)	-1.62	.11	-2.00	.06
Slow-speaking Context			10.4 (3.9)	10.0 (4.1)	0.97	.34	1.01	.32
Context Contrast	**By Subjects**	***t*(47)**	**-2.82**	0.42	Note: Significant contrasts are highlighted in bold			
		P	**.007**	.68				
	By Items	***t*(20)**	-1.75	0.41				
		p	.10	.69				

Table IV. Numbers of no-regressions go-past fixations launched in Experiment 10 (with standard deviations in brackets) **after removing items 1, 23, and 24 from analysis**. Also shown are results from 2-tailed paired sample *t*-tests examining the simple effects of *Context* and *Quoting Style*.

			Direct Speech	Indirect Speech	Quoting Style Contrast			
					By Subjects		By Items	
					t(47)	*p*	*t*(20)	*p*
Fast-speaking Context			9.5 (3.4)	9.8 (4.0)	-1.97	.06	-1.77	.09
Slow-speaking Context			10.0 (3.7)	9.5 (3.9)	1.86	.07	0.85	.41
Context Contrast	**By Subjects**	***t*(47)**	**-2.36**	0.84	Note: Significant contrasts are highlighted in bold			
		P	**.02**	.41				
	By Items	***t*(20)**	-1.63	0.12				
		p	.12	.91				

(ii) Average durations of fixations

The corresponding analyses on average **durations** of *go-past* (respectively *no-regressions go-past*) fixations revealed no reliable effects whatsoever, neither with the full set of items (all *p*s ≥ .15; maximum between-condition difference: 4 ms) nor when items 1, 23, and 24 were removed (all *p*s > .20; maximum between-condition difference: 5 ms).

Discussion

Taken together, these supplementary analyses suggest that the reading time effects reported in the paper are mainly due to modulations in the *numbers* of go-past fixations launched on the critical regions – average durations of those fixations were not affected much by the experimental conditions. It will be an interesting and important challenge for future research to pin down exactly what kind of eye-movement control mechanisms are involved in those pragmatically induced reading-rate modulations (e.g., whether they also involve changes in low-level word processing). However, such future research will almost certainly have to consider more measures and experimental conditions than we were able to include at present.

Appendix 7 Reading materials for Experiment 11

[D] = Direct speech; [I] = Indirect speech; [FF] = face focused; [FU] = face unfocused

The critical sentences (for acoustic analyses in oral reading) are in bold and italic text

Happy stories

1. Escaping from the hustle and bustle of the crowded city, couple Leona and Jason were enjoying an afternoon walk in the park. Leona revelled in this peaceful place as she saw baby ducks following their mother in the pond.
   [FF – D] She leaned on Jason with a satisfied smile: 'Aww, ***this is so relaxing after a busy week. It's almost magical***!'
   [FF- I] She leaned on Jason with a satisfied smile, saying that ***it was so relaxing after a busy week and that it was almost magical***.
   [FU – D] Leaning over, she gave Jason a big hug: 'Aww, ***this is so relaxing after a busy week. It's almost magical***!'
   [FU – I] Leaning over, she gave Jason a big hug, saying that ***it was so relaxing after a busy week and that it was almost magical***.

2. Derek's birthday was just a couple of days away and his girlfriend Ruby had arranged something special. A few days before, Ruby decided to surprise him with the plan that she kept hidden from him for so long.
   [FF- D] Grinning from ear to ear, she revealed: 'Look, darling, ***your birthday is coming up soon, so…I booked us tickets to London to visit the opera***!'
   [FF – I] Grinning from ear to ear, she revealed that since ***his birthday was coming up soon, she booked them tickets to London to visit the opera***.
   [FU – D] Stroking his cheek devotedly, she revealed: 'Look, darling, ***your birthday is coming up soon, so…I booked us tickets to London to visit the opera***!'
   [FU – I] Stroking his cheek devotedly, she revealed that since ***his birthday was coming up soon, she booked them tickets to London to visit the opera***.

3. In her dinner party, Vicky was a bit nervous because she had prepared some exotic recipes that she had never attempted before. Unexpectedly, all dishes went extremely popular and some guests even craved more after they finished.
   [FF – D] Vicky smiled jubilantly: 'Wow, ***I didn't expect that. I am so glad that you all enjoyed it***!'
   [FF – I] Vicky smiled jubilantly, saying that ***she did not expect that and she was so glad that they all enjoyed it***.
   [FU – D] Vicky felt rather relieved: 'Wow, ***I didn't expect that. I am so glad that you all enjoyed it***!'
   [FU – I] Vicky felt rather relieved, saying that ***she did not expect that and she was so glad that they all enjoyed it***.

4. In the small town of Oban lies a charming little pub. One evening, people were having a peaceful and relaxing night as usual when the famous writer Aileen burst in with exhilaration.
   [FF – D] She cheered with a big smile on her face: 'Guess what – ***my new novel has been accepted for publication! I'll buy everyone around***.'
   [FF – I] She cheered with a big smile on her face, announcing that ***her new novel had been accepted for publication and she would buy everyone a round***.
   [FU – D] Jumping up and down, she cheered: 'Guess what – ***my new novel has been accepted for publication! I'll buy everyone around***.'
   [FU – I] Jumping up and down, she cheered, announcing that ***her new novel had been accepted for publication and she would buy everyone a round***.

5. David was one of the best students in his school but he might not be able to go to university because of his poor family. He made a few applications, hoping that he might get some financial support or a scholarship and today he received a letter from Cambridge.
   [FF – D] Opening the envelope, he grinned widely: '***I got a full scholarship from University of Cambridge?! My parents will be very proud***.'
   [FF – I] Opening the envelope, he grinned widely, saying that ***he got a full scholarship from University of Cambridge and his parents would be very proud***.
   [FU – D] He thrilled to pieces as he opened the envelope, shouting: '***I got a full scholarship from University of Cambridge?! My parents will be very proud***!'
   [FU – I] He thrilled to pieces as he opened the envelope, shouting that ***he got a full scholarship from University of Cambridge and his parents would be very proud***.

6. The girls felt ravenous at 3am after hours of dancing and were desperately searching for a take away. After a few blocks, they finally gorged themselves in a chip shop and were now laughing at each other's faces as they ate.
[FF – D] With a hiccup, Hannah giggled: 'Oh, ***I love chips! They are just heavenly***!'
[FF – I] With a hiccup, Hannah giggled, saying that ***she loved chips and they were just heavenly***.
[FU – D] With a hiccup, Hannah stroked her tummy contentedly: 'Oh, ***I love chips! They are just heavenly***!'
[FU – I] With a hiccup, Hannah stroked her tummy contentedly, saying that ***she loved chips and they were just heavenly***.

7. It was the last day of John Lewis's summer sale when two girls Helen and Abbey were browsing around. Helen was very fond of a black skirt and tried it on straightaway when she found it was heavily discounted.
[FF – D] She was grinning widely: 'Hey! ***How do I look? It is exactly what I always wanted, and now it is even more than 50% off***!'
[FF – I] Grinning widely, she asked ***how she looked and said that it was exactly what she always wanted, and that now it was even more than 50% off***.
[FU – D] Jumping up and down, she showed it to Abbey: 'Hey! ***How do I look? It is exactly what I always wanted, and now it is even more than 50% off***!'
[FU – I] Jumping up and down, she showed it to Abbey, asking ***how she looked and said that it was exactly what she always wanted, and that now it was even more than 50% off***.

8. Business woman Stella had a rather sunny and cheerful day because she was just promoted to the department manager. When she got home, she was surprised that her husband had already prepared a candle-light dinner to congratulate her.
[FF – D] She hugged him with a contented smile: 'Thank you, my darling. ***This is very sweet of you***!'
[FF – I] She hugged him with a contented smile, thank him and said that ***it was very sweet of him***.
[FU – D] Contentedly, she gave him a big hug: 'Thank you, my darling. ***This is very sweet of you***!'
[FU – I] Contentedly, she gave him a big hug to thank him, saying that ***it was very sweet of him***.

9. After 2 weeks of international conferences, Lizzi finally arrived home and her puppy dashed into her arms, licking her face ecstatically. Lizzi's exhaustion was instantly wiped away by such a heart-warming welcome from her little dog.
[FF – D] She giggled: 'Aww darling. ***I missed you too! I dreamt of you every single night in the past two weeks***.'
[FF – I] She giggled, saying that ***she missed her too and she dreamt about her every single night in the past two weeks***.
[FU – D] She stroked him devotedly: 'Aww darling. ***I missed you too! I dreamt of you every single night in the past two weeks***.'
[FU – I] She stroked him devotedly, saying that ***she missed her too and she dreamt about her every single night in the past two weeks***.

10. It was 25th Dec and school girl Melissa got up at 7am because she was eager to find out what she had received from Santa Claus. A feeling of pure joy came over her as she opened the big pink box – a Barbie doll called 'Emma'!
[FF – D] She ran upstairs with a big smile on her face: 'Mummy, ***Santa Claus has sent me a sister and her name is Emma***!'
[FF – I] She ran upstairs with a big smile on her face, telling her mum that ***Santa Claus had sent her a sister and her name was Emma***.
[FU – D] She ran upstairs, shouting joyfully: 'Mummy, ***Santa Claus has sent me a sister and her name is Emma***!'
[FU – I] She ran upstairs, shouting joyfully to her mum that ***Santa Claus had sent her a sister and her name was Emma***.

11. PhD student Ella had just finished her viva and was now wandering nervously outside the exam room, waiting for the final decision. Her friends and colleagues had gathered to support her until the examiners came out with a round of applause.
[FF – D] Over the moon, Ella was grinning from ear to ear: 'Oh, ***I would not have gone through this without you, my friends! Now, it is time for us to celebrate***!'
[FF – I] Over the moon, Ella was grinning from ear to ear, saying that ***she would not have gone that far without them and now it was time for them to celebrate***.
[FU – D] Thrilled to pieces, Ella hugged her friends joyously: 'Oh, ***I would not have gone through this without you, my friends! Now, it is time for us to celebrate***!'
[FU – I] Thrilled to pieces, Ella hugged her friends joyously, saying that ***she would not have gone that far without them and now it was time for them to celebrate***.

12. It was approaching the end of the term; first-year student Katie was withdrawing her essay mark in the lab. Although she had worked very hard, she was slightly worried because many students had received a C.
[FF – D] She smiled jubilantly as she saw an A-: 'Oh thank god. ***I was very worried. I am so glad that my efforts paid off***!'
[FF – I] She smiled jubilantly as she saw an A-, saying that ***she had been very worried and she was so glad that her efforts had paid off***.
[FU – D] She felt relieved as she saw an A-: 'Oh thank god. ***I was very worried. I am so glad that my efforts paid off***!'
[FU – I] She felt relieved as she saw an A-, saying that ***she had been very worried and she was so glad that her efforts had paid off***.

Surprised Stories (positive)

13. Alex was a world-renowned art collector who was specialized in Italian renaissance paintings. He bought a second-hand painting box for his granddaughter and was shocked when he discovered the inscription: '1472, Leonardo da Vinci' at the inner edge of the box.
[FF – D] His eyes became wide open: 'Oh my goodness, ***I can't believe this is da Vinci's painting box***!'
[FF – I] Eyes wide open, he said that ***he couldn't believe it was da Vinci's painting box***.
[FU – D] He was completely gobsmacked: 'Oh my goodness, ***I can't believe this is da Vinci's painting box***!'
[FU – I] He was completely gobsmacked, saying that ***he couldn't believe it was da Vinci's painting box***.

14. Conservation students Sarah and Brian drifted along the coast in their sea kayak, spotting local birds and wildlife. As they rounded a cliff, they came across a breath-taking Sanford's Sea-eagle which was only 5 meters away from them.
[FF – D] Brian's mouth fell open in surprise: 'Bloody hell! ***This is a Sanford's Sea-eagle! It has never been found in Britain before***!'
[FF – I] Brian's mouth fell open in surprise and he said that ***it was a Sanford's Sea-eagle and it had never been found in Britain before***.
[FU – D] Holding his breath, Brian was shocked: 'Bloody hell! ***This is a Sanford's Sea-eagle! It has never been found in Britain before***!'
[FU – I] Holding his breath, Brian was shocked, saying that ***it was a Sanford's Sea-eagle and it had never been found in Britain before***.

15. In Beijing, tourists Alison and Nick were going to a big silk market because they had never tried silk clothes before. When they entered the market, they found themselves overwhelmed by a wide range of fine fabric on offer.
[FF – D] Alison touched one and her eyes were instantly widened: 'Nick! ***The silk feels so incredibly smooth***!'
[FF – I] As Alison touched one, her eyes were instantly widened and she told Nick that ***the silk felt so incredibly smooth***.
[FU – D] Alison's voice was filled with surprise: 'Nick! ***The silk feels so incredibly smooth***!'
[FU – I] Alison's voice was filled with surprise as she told Nick that ***the silk felt so incredibly smooth***.

16. Dave had been a lucky postman because he had won a fortune in the lottery and soon purchased an expensive Ferrari 328GTS Targa. Today in the car park, his colleague Ralph was utterly shocked by Dave's new possession.
[FF – D] His mouth fell open in surprise: 'Wow! ***Is that your new car!? It looks really stylish***!'
[FF – I] His mouth fell open in surprise as he asked Dave ***whether that was his new car, adding that it looked really stylish***.
[FU – D] His voice filled with envy: 'Wow! ***Is that your new car!? It looks really stylish***!'
[FU – I] His voice filled with envy, asking Dave ***whether that was his new car, adding that it looked really stylish***.

17. Peter was upset because his dad - his best mate - did not give him any presents for his 18th birthday today. He went to the garage to fetch his soccer ball and was rooted to the spot as he saw a brand new Toyota inscribed with 'Peter's 18th, love dad'.
[FF – D] His eyes became wide open: 'Good lord! ***I'm not dreaming, am I? It looks absolutely amazing***!'
[FF – I] Eyes wide open, he asked ***if he was dreaming, adding that it looked absolutely amazing***.
[FU – D] He was completely gobsmacked: 'Good lord! ***I'm not dreaming, am I? It looks absolutely amazing***!'
[FU – I] He was completely gobsmacked, asking ***if he was dreaming, adding that it looked absolutely amazing***.

18. In MRC biomedical laboratory, Dr. Parkinson was browsing the microscopic photos taken in their recent experiments. As he saw the effects of peptide-24 on amyloid β proteins on the infected brain tissues, he was utterly astonished.
[FF – D] His eyes were widened: 'A… ***A significant shrinkage! This could be a potential cure for Alzheimer's disease***!'
[FF – I] Eyes widened, he said that there was ***a significant shrinkage and that it could be a potential cure for Alzheimer's disease***.
[FU – D] He breathed rapidly: 'A… ***A significant shrinkage! This could be a potential cure for Alzheimer's disease***!'
[FU – I] He breathed rapidly, saying that there was ***a significant shrinkage and that it could be a potential cure for Alzheimer's disease***.

Surprise stories (negative)

19. At the beach, William and Phil were enjoying the sunshine and a fresh sea breeze. Suddenly, William noticed a piece of wood floating on the water with something lying on it that looked like an unconscious child.
[FF – D] Eyes wide open, he gasped: 'Phil, quick! ***We need to call for help because there is a child in trouble***!'
[FF – I] Eyes wide open, he gasped and told Phil that ***they needed to call for help because there was a child in trouble***.
[FU – D] He was numb with shock: 'Phil, quick! ***We need to call for help because there is a child in trouble***!'
[FU – I] He was numb with shock, telling Phil that ***they needed to call for help because there was a child in trouble***.

20. At a peaceful house in Dumbarton Mrs. Harris was preparing lunch for her children. The phone rang and Mrs. Harris was stunned by what she heard on the phone.
[FF – D] Her face looked rather frightened and she said in horror: 'Oh no! ***My husband was involved in a collision with a heavy goods vehicle and is now in hospital***!'
[FF – I] Her face looked rather frightened and she said in horror that ***her husband was involved in a collision with a heavy goods vehicle and was now in hospital***.
[FU – D] She was dazed at the scene, muttering: 'Oh no! ***My husband was involved in a collision with a heavy goods vehicle and is now in hospital***!'
[FU – I] She was dazed at the scene, muttering that ***her husband was involved in a collision with a heavy goods vehicle and was now in hospital***.

21. In the library, university student Suzanne was having a snack break from busy exam revision. After two packs of Milk Chocolate-Covered Oreos, she instantly gasped when she read the back of the package.
[FF – D] Her eyes nearly came out of her sockets: 'Oh no!! ***What have I done! I just had a 2000-calorie snack***!'
[FF – I] Her eyes nearly came out of her sockets while she asked herself ***what she had done, adding that she just had a 2000-calorie snack***.
[FU – D] She was grabbing her hair: 'Oh no!! ***What have I done! I just had a 2000-calorie snack***!'
[FU – I] She was grabbing her hair while she asked herself ***what she had done, adding that she just had a 2000-calorie snack***.

22. There had been a couple of complaints about the food in the 1-Michelin star restaurant Alain Ducasse this evening. The restaurant manager came into the kitchen for a look and was instantly astounded by its unacceptable state.
[FF – D] His face looked rather frightened: '***This cannot be true! The fish is burnt, the vegetables overcooked, and the soup is actually stone cold...***'
[FF – I] His face looked rather frightened as he said that ***it could not be true that the fish was burnt, the vegetables overcooked, and the soup was actually stone cold***.
[FU – D] He was grabbing his hair: '***This cannot be true! The fish is burnt, the vegetables overcooked, and the soup is actually stone cold...***'
[FU – I] He was grabbing his hair, saying that ***it could not be true that the fish was burnt, the vegetables overcooked, and the soup was actually stone cold***.

23. It was one week before Christmas and Elaine was still struggling to finish all her work before her holiday. Her glimmer of hope for a relaxing holiday was now mercilessly destroyed by yet another huge pile of files on her desk.

[FF – D] Her eyes nearly came out of her sockets: '***I don't believe this. There seems to be no end to it in sight...***'
[FF – I] Her eyes nearly came out of her sockets while she said ***she couldn't believe it and that there seemed to be no end to it in sight***.
[FU – D] Looking at the pile, Elaine was numb with shock: '***I don't believe this. There seems to be no end to it in sight...***'
[FU – I] Looking at the pile, Elaine was numb with shock, saying that ***she couldn't believe it and that there seemed to be no end to it in sight***.

24. English fan John's blissful happiness from Steven Gerrard's leading goal against USA did not last long. Shockingly, the English goalkeeper Robert Green inexplicably allowed USA's soft, skidding shot to escape his grasp and roll into the net.

    [FF – D] Eyes wide open, he gasped: '***I don't believe this. Green, you must have butter fingers***!'
    [FF – I] Eyes wide open, he gasped and said that ***he did not believe what had happened and that Green must have butter fingers***.
    [FU – D] John was dazed at the scene, muttering: '***I don't believe this. Green, you must have butter fingers***!'
    [FU – I] John was dazed at the scene, muttering that ***he did not believe what had happened and that Green must have butter fingers***.

Angry stories

25. First-year university student Britney found that the weather in Glasgow was one of the worst in Britain. After a heavy snow in the afternoon, she fell on her elbow and found that her new, expensive jacket was dirty and scratched.
    [FF – D] She complained with an angry face: '***I really hate the winter! It's always dark and the roads are too slippery***.'
    [FF – I] She complained with an angry face that ***she really hated the winter because it was always dark and the roads were too slippery***.
    [FU – D] She complained in an angry voice: '***I really hate the winter! It's always dark and the roads are too slippery***.'
    [FU – I] She complained in an angry voice that ***she really hated the winter because it was always dark and the roads were too slippery***.

26. Universities in England were preparing to raise tuition fees following the government's cuts of 40%.Simon, among thousands of students, was angrily protesting outside Parliament today. When he was interviewed, he appeared particularly indignant.
    [FF – D] He frowned disapprovingly: 'You know what, ***the MPs have broken their promises, peddle lies about fairness and we will expose their betrayal***!'
    [FF – I] He frowned disapprovingly, saying that ***the MPs had broken their promises, peddled lies about fairness and that the students would expose their betrayal***.
    [FU – D] He aired his grievances: 'You know what, ***the MPs have broken their promises, peddle lies about fairness and we will expose their betrayal***!'
    [FU – I] He aired his grievances, saying that ***the MPs had broken their promises, peddled lies about fairness and that the students would expose their betrayal***.

27. The nursery was filled with lots of young children, screaming for attention from the busy childminder. One boy in particular had been told off many times for behaving disgracefully all day and yet he was still throwing sponge balls at the poor lady.
    [FF – D] With an annoyed frown, she stressed to him: 'Listen. ***You have to be really good and not make a mess for the next few hours***.'
    [FF – I] With an annoyed frown, she stressed to him that ***he had to be really good and not make a mess for the next few hours***.
    [FU – D] Pointing at the boy's nose, she exploded: 'Listen. ***You have to be really good and not make a mess for the next few hours***.'
    [FU – I] Pointing at the boy's nose, she exploded, saying that ***he had to be really good and not make a mess for the next few hours***.

28. A trendy night club had been recently opened near Marvin's flat and it proved instantly popular. Every night, Marvin was terribly disturbed at sleep by the thudding noise and this morning he was venting his anger to his neighbour.
    [FF – D] His face filled with resentment: '***This club is ridiculous! Every time I have to work early hours, they are having a bash***!'
    [FF – I] His face was filled with resentment while he said that ***the club was ridiculous because every time he had to work early hours, they were having a bash***.
    [FU – D] He was blustering: '***This club is ridiculous! Every time I have to work early hours, they are having a bash***!'
    [FU – I] He was blustering that ***the club was ridiculous because every time he had to work early hours, they were having a bash***.

29. Jenny and Irvin were taking a holiday trip to Barcelona and they were very glad that they had booked a very cheap flight. However, it turned out that they had to pay extra money for virtually every service on plane, which even included £3 for the toilet.
[FF – D] Her face filled with rage: 'What!? ***I have to pay for a pee? How ridiculous! This flight is an absolute rip-off***!'
[FF – I] Her face filled with rage while she said that ***it was ridiculous for her to pay for a pee, concluding that the flight was an absolute rip-off***.
[FU – D] She became increasingly cross, shouting: 'What!? ***I have to pay for a pee? How ridiculous! This flight is an absolute rip-off***!'
[FU – I] She became increasingly cross, shouting that ***it was ridiculous for her to pay for a pee, concluding that the flight was an absolute rip-off***.

30. James was annoyed when an important mail he had asked to be sent urgently was returned to the company in the post after two weeks. He was in a foul mood and stormed through to his assistant's office, not bothering to knock on the door.
[FF – D] Banging the package on the desk, James gnashed his teeth: '***If you don't want to be fired, make sure it gets to the right person this time***!'
[FF – I] Banging the package on the desk, James gnashed his teeth, saying that ***if the assistant did not want to be fired, he should make sure it got to the right person this time***.
[FU – D] Banging the package on the desk, James shouted furiously: '***If you don't want to be fired, make sure it gets to the right person this time***!'
[FU – I] Banging the package on the desk, James shouted furiously that ***if the assistant did not want to be fired, he should make sure it got to the right person this time***.

31. Millionaire Joseph was addicted to betting on horses and he had always been dreaming of a big win. Tipped by a so-called 'insider', he recently placed an enormous bet, but shockingly, the horse had lost.
[FF – D] His face was filled with resentment while he was on the phone: '***Where did your bloody information come from!? That was a huge amount of money – almost one million pounds***!'
[FF – I] His face was filled with resentment while he was on the phone, asking ***where his bloody information came from, adding that it was a huge amount of money – almost one million pounds***.
[FU – D] Angry with his informant, Joseph shouted furiously on the phone: '***Where did your bloody information come from!? That was a huge amount of money – almost one million pounds***!'
[FU – I] Angry with his informant, Joseph shouted furiously on the phone asking ***where his bloody information came from, adding that it was a huge amount of money – almost one million pounds***.

32. Leanne's son had promised her that he would tidy his room when he got home from school. She was furious to find him watching a film, his speakers blaring and his room still a mess.
[FF – D] She lunged to turn the TV off and gnashed her teeth: '***That's it! You're going to be grounded until this room is sparkling***.'
[FF – I] She lunged to turn the TV off and gnashed her teeth, saying ***that was it and he was going to be grounded until that room was sparkling***.
[FU – D] She lunged to turn the TV off and exploded: '***That's it! You're going to be grounded until this room is sparkling***.'
[FU – I] She lunged to turn the TV off and exploded, saying ***that was it and he was going to be grounded until that room was sparkling***.

33. It was Sunday afternoon when Heather finally came back home after visiting her parents. She was outraged to find her boyfriend Ben sleeping like a pig in his messy bedroom where she couldn't even find a space to stand.
[FF – D] She dragged Ben off the bed and roared with an angry face: '***I can't believe how messy you are! The first thing I found on the floor were your dirty pyjamas***!'
[FF – I] She dragged Ben off the bed and roared with an angry face, saying that ***she could not believe how messy he was and that the first thing she found on the floor were his dirty pyjamas***.
[FU – D] She dragged Ben off the bed and roared in his face: '***I can't believe how messy you are! The first thing I found on the floor were your dirty pyjamas***!'
[FU – I] She dragged Ben off the bed and roared in his face, saying that ***she could not believe how messy he was and that the first thing she found on the floor were his dirty pyjamas***.

34. During a coffee break, university teachers Robert and Isabel met in the common room. Robert was irritated with the mess in the kitchen and according to a post-it, his mug had even been 'borrowed' by someone else.
[FF – D] His face was filled with rage: '***How brazen are these people! They didn't want to wash their dirty mugs so they just 'borrowed' mine!***'
[FF – I] His face was filled with rage as he said that ***these people were very brazen because when they didn't want to wash their dirty mugs, they just 'borrowed' his***.
[FU – D] He aired his grievances to Isabel: '***How brazen are these people! They didn't want to wash their dirty mugs so they just 'borrowed' mine***!'
[FU – I] He aired his grievances to Isabel saying that ***these people were very brazen because when they didn't want to wash their dirty mugs, they just 'borrowed' his***.

35. Shauna was a young, talented actress and singer but she had been unemployed for months. She was now desperate for a part in a new musical project but her boyfriend was ill and blamed her for not looking after him.
[FF – D] With an annoyed frown, Shauna stressed to him: 'Don't be ridiculous! ***I am doing this because without this job I can't even pay for the electricity bill***!'
[FF – I] With an annoyed frown, Shauna stressed to him that he was ridiculous and ***she was doing that because without that job she couldn't even pay for the electricity bill***.
[FU – D] Shauna was angry with his attitude, blustering: 'Don't be ridiculous! ***I am doing this because without this job I can't even pay for the electricity bill***!'
[FU – I] Shauna was angry with his attitude, blustering that he was ridiculous and ***she was doing that because without that job she couldn't even pay for the electricity bill***.

36. One morning, six-year-old Sara was taking a lot of time to get ready for school and the school bus was due to arrive soon. Although her dad had been reminding her for half hour, she was still in her pyjamas.
[FF – D] Her dad frowned disapprovingly: 'Sara! ***If you are not ready within five minutes, I'll have to drive you again***!'
[FF – I] Her dad frowned disapprovingly, saying that ***if she were not ready within five minutes, he would have to drive her again***.
[FU – D] Her dad became increasingly cross, shouting: 'Sara! ***If you are not ready within five minutes, I'll have to drive you again***!'
[FU – I] Her dad became increasingly cross, shouting that ***if she were not ready within five minutes, he would have to drive her again***.

References

Abramson, M. (2007). The written voice: Implicit memory effects of voice characteristics following silent reading and auditory presentation. *Perceptual and Motor Skills, 105*(3), 1171-1186.

Abramson, M., & Goldinger, S. D. (1997). What the reader's eye tells the mind's ear: Silent reading activates inner speech. *Perception & Psychophysics, 59*(7), 1059-1068.

Alexander, J. D., & Nygaard, L. C. (2008). Reading voices and hearing text: Talker-specific auditory imagery in reading. *Journal of Experimental Psychology-Human Perception and Performance, 34*(2), 446-459.

Anderson, J. R. (1974). Verbatim and propositional representation of sentences in immediate and long-term memory. *Journal of Verbal Learning and Verbal Behavior, 13*(2), 149-162.

Ashby, J., & Clifton, C. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition, 96*(3), B89-B100.

Ashby, J., & Rayner, K. (2004). Representing syllable information during silent reading: Evidence from eye movements. *Language and Cognitive Processes, 19*(3), 391-426.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614-636.

Banziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication, 46*(3-4), 252-267.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577-609.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617-645.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B-Biological Sciences, 364*(1521), 1281-1289.

Barsalou, L. W. (2010). Grounded Cognition: Past, Present, and Future. *Topics in Cognitive Science, 2*(4), 716-724.

Beattie, G., Cutler, A., & Pearson, M. (1982). 'Why is Mrs Thatcher interrupted so often?'. *Nature, 300*, 744-747.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature, 403*(6767), 309-312.

Blyth, C., Recktenwald, S., & Wang, J. (1990). I'm like, 'Say what?!': A new quotative in American oral narrative. *American Speech, 65*(3), 215-227.

Boersma, P., & Weenink, D. (2010). Praat: doing phonetics by computer. Version 5.1.43, retrieved 4 August 2010 from http://www.praat.org/.

Bohan, J., Sanford, A. J., Cochrane, S., & Sanford, A. J. S. (2008). *Direct and indirect speech modulates depth of processing.* Paper presented at the 14th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP), Cambridge, UK.

Bolinger, D. (1964). *Intonation as a universal*. Paper presented at the Proceedings of the 5th Congress of Phonetics.

Borreggine, K. L., & Kaschak, M. P. (2006). The action-sentence compatibility effect: It's all in the timing. *Cognitive Science, 30*(6), 1097-1112.

Brown, C. A., & Bacon, S. P. (2010). Fundamental frequency and speech intelligibility in background noise. *Hearing Research, 266*(1-2), 52-59.

Bubic, A., von Cramon, D. Y., Jacobsen, T., Schroger, E., & Schubotz, R. I. (2009). Violation of Expectation: Neural Correlates Reflect Bases of Prediction. *Journal of Cognitive Neuroscience, 21*(1), 155-168.

Bulut, M., & Narayanan, S. (2008). On the robustness of overall F0-only modifications to the perception of emotions in speech. *Journal of the Acoustical Society of America, 123*(6), 4547-4558.

Bunzeck, N., Wuestenberg, T., Lutz, K., Heinze, H. J., & Jancke, L. (2005). Scanning silence: Mental imagery of complex sounds. *Neuroimage, 26*(4), 1119-1127.

Busso, C., Lee, S., & Narayanan, S. (2009). Analysis of Emotionally Salient Aspects of Fundamental Frequency for Emotion Detection. *Ieee Transactions on Audio Speech and Language Processing, 17*(4), 582-596.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., McGuire, P. K., & David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science, 276*(5312), 593-596.

Chafe, W. L. (Ed.). (1982). *Integration and involvement in speaking, writing, and oral literature*: Norwood, NJ: Ablex.

Chomsky, N. (1957). *Syntactic Structures*: The Hague: Mouton.

Clark, H. H., & Gerrig, R. J. (1990). Quotations as Demonstrations. *Language, 66*(4), 764-805.

Cohen, A., Collier, R., & Thart, J. (1982). Declination: Construct or intrinsic feature of speech pitch? [Article]. *Phonetica, 39*(4-5), 254-273.

Cohen, A., & t' Hart, J. (1967). On the anatomy of intonation. *Lingua, 19*, 177-192.

Connell, L. (2007). Representing object colour in language comprehension. *Cognition, 102*(3), 476-485.

Coulmas, F. (Ed.). (1986). *Direct and Indirect Speech*. Berlin; New York; Amsterdam: Mouton de Gruyter.

Cox, R. M., Alexander, G. C., Taylor, I. M., & Gray, G. A. (1997). The contour test of loudness perception. *Ear and Hearing, 18*(5), 388-400.

Cukor-Avila, P. (2002). She say, she go, she be like: Verbs of quotation over time in African American Vernacular English. *American Speech, 77*(1), 3-31.

Daniel, C. (2009). Embodiment of abstract concepts: Good and bad in right- and left-handers. *Journal of Experimental Psychology: General, 138*(3), 351-367.

Danly, M., Cooper, W. E., & Shapiro, B. (1983). Fundamental-frequency, language processing, and linguistic structure in Wernicke's aphasia. *Brain and Language, 19*(1), 1-24.

Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research, 229*(1-2), 132-147.

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *NeuroImage, 55*(2), 705-712.

Dzhelyova, M. P., Ellison, A., & Atkinson, A. P. (2011). Event-related repetitive TMS reveals distinct, critical roles for right OFA and bilateral posterior STS in judging the sex and trustworthiness of faces. *Journal of Cognitive Neuroscience, 23*(10), 2782-2796.

Enrici, I., Adenzato, M., Cappa, S., Bara, B. G., & Tettamanti, M. (2011). Intention processing in communication: A common brain network for language and gestures. *Journal of Cognitive Neuroscience, 23*(9), 2415-2431.

Ferrara, K., & Bell, B. (1995). Sociolinguistic variation and discourse function of constructed dialogue introducers - the case of be + like. *American Speech, 70*(3), 265-290.

Fincher-Kiefer, R. (2001). Perceptual components of situation models. *Memory & Cognition, 29*(2), 336-343.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology, 61*(6), 825-850.

Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*(3), 558-565.

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology, 61*(6), 905-919.

Goodell, E. W., & Sachs, J. (1992). Direct and indirect speech in English-speaking children's retold narratives. *Discourse Processes, 15*(4), 395-422.

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: Brain responses to angry prosody in meaningless speech. *Nature Neuroscience, 8*(2), 145-146.

Hardin, J. W., & Hilbe, J. M. (2003). *Generalized Estimating Equations*. London: Chapman & Hall/CRC.

Harnad, S. (1990). The symbol grounding problem. *Physica D, 42*(1-3), 335-346.

Havas, D. A., Glenberg, A. M., Gutowski, K. A., Lucarelli, M. J., & Davidson, R. J. (2010). Cosmetic use of Botulinum Toxin-A affects processing of emotional language. *Psychological Science, 21*(7), 895-900.

Hird, K., & Kirsner, K. (2002). The relationship between prosody and breathing in spontaneous discourse. *Brain and Language, 80*(3), 536-555.

Hjelmquist, E., & Gidlund, Å. (1985). Free recall of conversations. *Text - Interdisciplinary Journal for the Study of Discourse, 5*(3), 169 - 186.

Holmes, K. J., & Lourenco, S. F. (2011). Common spatial organization of number and emotional expression: A mental magnitude line. *Brain and Cognition, 77*(2), 315-323.

Holt, E. (1996). Reporting on talk: The use of direct reported speech in conversation. *Research on Language and Social Interaction, 29*(3), 219-245.

Hooker, C. I., Paller, K. A., Gitelman, D. R., Parrish, T. B., Mesulam, M. M., & Reber, P. J. (2003). Brain networks for analyzing eye gaze. *Cognitive Brain Research, 17*(2), 406-418.

Huettel, S. A., Mack, P. B., & McCarthy, G. (2003). Brain systems associated with violations of expectation. *Society for Neuroscience Abstract Viewer and Itinerary Planner, 2003*, Abstract No. 343.347.

Jancke, L., & Shah, N. J. (2004). 'Hearing' syllables by 'seeing' visual stimuli. *European Journal of Neuroscience, 19*(9), 2603-2608.

Jarvella, R. J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior, 10*(4), 409-416.

Johnsrude, I. S., Penhune, V. B., & Zatorre, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain, 123*, 155-163.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*(11), 4302-4311.

Kaschak, M. P., & Borreggine, K. L. (2008). Temporal dynamics of the action-sentence compatibility effect. *Quarterly Journal of Experimental Psychology, 61*(6), 883-895.

Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language, 43*(3), 508-529.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition, 94*(3), B79-B89.

Kaup, B., Ludtke, J., & Zwaan, R. A. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics, 38*(7), 1033-1050.

Kellenbach, M. L., Brett, M., & Patterson, K. (2001). Large, colorful, or noisy? Attribute- and modality-specific activations during retrieval of perceptual attribute knowledge. *Cognitive Affective and Behavioral Neuroscience, 1*(3), 207-221.

Kent, R. D., & Read, C. (1992). *The acoustic analysis of speech*. San Diego: Singular Publishing Group.

Kiefer, M., & Pulvermüller, F. (in press). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*.

Kiefer, M., Sim, E. J., Herrnberger, B., Grothe, J., & Hoenig, K. (2008). The sound of concepts: four markers for a link between auditory and conceptual brain systems. *Journal of Neuroscience, 28*(47), 12224-12230.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, MA: Cambridge University Press.

Kintsch, W., & Vandijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394.

Klein, M. E., & Zatorre, R. J. (2011). A role for the right superior temporal sulcus in categorical perception of musical chords. *Neuropsychologia, 49*(5), 878-887.
Kosslyn, S. M., & Matt, A. M. C. (1977). If you speak slowly, do people read your prose slowly - person-perticular speech recoding during reading. *Bulletin of the Psychonomic Society, 9*(4), 250-252.
Kring, A. M., & Gordon, A. H. (1998). Sex differences in emotion: Expression, experience, and physiology. *Journal of Personality and Social Psychology, 74*(3), 686-703.
Kurby, C. A., Magliano, J. P., & Rapp, D. N. (2009). Those voices in your head: Activation of auditory images during reading. *Cognition, 112*(3), 457-461.
Labov, W. (1972). The transformation of experience in narrative syntax. *Language in the inner city* (pp. 354-396): Philadelphia: University of Pennsylvania Press.
Leech, G. N., & Short, M. H. (1981). *Style in fiction: A linguistic introduction to English fictional prose*. London: Longman.
Lehrer, A. (1989). Remembering and representing prose - quoted speech as a data source. *Discourse Processes, 12*(1), 105-125.
Lynott, D., & Connell, L. (2009). Modality exclusivity norms for 423 object properties. *Behavioral Research Methods, 41*(2), 558-564.
MacSweeney, M., Amaro, E., Calvert, G. A., Campbell, R., David, A. S., McGuire, P., & Brammer, M. J. (2000). Silent speechreading in the absence of scanner noise: an event-related fMRI study. [Article]. *NeuroReport, 11*(8), 1729-1733.
Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M., & Frank, M. (Eds.). (2008). *Facial Expressions of Emotion*. New York; London: the Guilford Press.
Menn, L., & Boyce, S. (1982). Fundamental frequency and discourse structure. *Language and Speech, 25*(4), 341-383.
Mitchell, R. L. (2006). How does the brain mediate interpretation of incongruent auditory emotions? The neural response to prosody in the presence of conflicting lexico-semantic cues. *European Journal of Neuroscience, 24*(12), 3611-3618.
Mitchell, R. L. C., Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. R. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia, 41*(10), 1410-1421.
Morrison, R., & Rayner, K. (1981). Saccade size in reading depends upon character spaces and not visual angle. *Attention, Perception, & Psychophysics, 30*(4), 395-396.
Niedenthal, P. M., Mermillod, M., Maringer, M., & Hess, U. (2010). The Simulation of Smiles (SIMS) model: Embodied simulation and the meaning of facial expression. *Behav Brain Sci, 33*(6), 417-433; discussion 433-480.
Ogawa, S., & Lee, T. M. (1990). Magnetic-resonance-imaging of blood-vessels at high fields: in vivo and in vitro measurements and image stimulation. *Magnetic Resonance in Medicine, 16*(1), 9-18.
Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic-resonance-imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America, 87*(24), 9868-9872.
Ogawa, S., Lee, T. M., Nayak, A. S., & Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic-resonance image of rodent brain at high magnetic-fields. *Magnetic Resonance in Medicine, 14*(1), 68-78.
Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science, 14*(2), 119-124.
Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal Representations of Perceived Emotions in the Human Brain. *Journal of Neuroscience, 30*(30), 10127-10134.
Pell, M. D. (2001). Influence of emotion and focus location of prosody in matched statements and questions. *Journal of the Acoustical Society of America, 109*(4), 1668-1680.
Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *Neuroimage, 56*(4), 2356-2363.

Pitt, M. A., & Crowder, R. G. (1992). The role of spectral and dynamic cues in imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance, 18*(3), 728-738.

Pulvermuller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience, 6*(7), 576-582.

Pylyshyn, Z. W. (1984). *Computation and cognition: Toward a foundation for cognitive science*: Cambridge, MA: MIT Press.

Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: looking at things that aren't there anymore. *Cognition, 76*(3), 269-295.

Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science, 27*(5), 767-780.

Robinson, M. D., & Johnson, J. T. (1997). Is it emotion or is it stress? Gender stereotypes and the perception of subjective experience. *Sex Roles, 36*(3-4), 235-258.

Roy, C. S., & Sherrington, C. S. (1890). On the regulation of the blood-supply of the brain. *The Journal of physiology, 11*(1-2), 85-158.117.

Russell, J. A., Bachorowski, J. A., & Fernandez-Dols, J. M. (2003). Facial and vocal expressions of emotion. *Annual Review of Psychology, 54*, 329-349.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics, 2*(9), 437-442.

Schiffrin, D. (1981). Tense variation in narrative. *Language, 57*(1), 45-62.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain, 123*, 2400-2406.

Searle, J. R. (1980). Minds, Brains, and Programs. *Behavioral and Brain Sciences, 3*(3), 417-425.

Shergill, S. S., Bullmore, E. T., Brammer, M. J., Williams, S. C. R., Murray, R. M., & McGuire, P. K. (2001). A functional study of auditory verbal imagery. *Psychological Medicine, 31*(2), 241-253.

Short, M., Semino, E., & Wynne, M. (2002). Revisiting the notion of faithfulness in discourse presentation using a corpus approach. *Language and Literature, 11*(4), 325 - 355.

Sokolov, A. A., Erb, M., Gharabaghi, A., Grodd, W., Tatagiba, M. S., & Pavlova, M. A. (2011). Biological motion processing: The left cerebellum communicates with the right superior temporal sulcus. *Neuroimage*.

Spitsyna, G., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. S. (2006). Converging language streams in the human temporal lobe. *Journal of Neuroscience, 26*(28), 7328-7336.

Stafford, L., Burggraf, C. S., & Sharkey, W. F. (1987). Conversational memory - the effects of time, recall, mode, and memroy expectancies on remembrances of natural conversations. *Human Communication Research, 14*(2), 203-229.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12*(2), 153-156.

Sturt, P., Sanford, A. J., Stewart, A., & Dawydiak, E. (2004). Linguistic focus and good-enough representations: An application of the change-detection paradigm. *Psychonomic Bulletin & Review, 11*(5), 882-888.

Tagliamonte, S., & D'Arcy, A. (2004). He's like, she's like: The quotative system in Canadian youth. *Journal of Sociolinguistics, 8*(4), 493-514.

Tannen, D. (1982). Oral and Literate Strategies in Spoken and Written Narratives. *Language, 58*(1), 1-21.

Tannen, D. (1986). Introducing constructed dialogue in Greek and American conversational and literary narrative. In F. Coulmas (Ed.), *Direct and Indirect Speech*. Berlin; New York; Amsterdam: Mouton de Gruyter.

Tannen, D. (1989). 'Oh talking voice that is so sweet': Constructing dialogue in conversation. In D. Tannen (Ed.), *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.

Taylor, L. J., & Zwaan, R. A. (2008). Motor resonance and linguistic focus. *Quarterly Journal of Experimental Psychology, 61*(6), 896-904.

Thomson, A. J., & Martinet, A. V. (1986). *A practical English grammar* (4th ed.). Oxford: Oxford University Press.

Tree, J. E. F., & Tomlinson, J. M. (2008). The rise of like in spontaneous quotations. *Discourse Processes, 45*(1), 85-102.

Umeda, N. (1981). Influence of segmental factors on fundamental-frequency in fluent speech. *Journal of the Acoustical Society of America, 70*(2), 350-355.

Vaissiere, J. (Ed.). (1983). *Language-independent prosodic features*. New York: Springer-Verlag.

Van Essen, D. C. (2002). Windows on the brain: the emerging role of atlases and databases in neuroscience. *Current Opinion in Neurobiology, 12*(5), 574-579.

Van Essen, D. C., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., & Anderson, C. H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *Journal of the American Medical Informatics Association, 8*(5), 443-459.

Van Essen, D. C., Harwell, J., Hanlon, D., & Dickson, J. (Eds.). (2005). *Surface-Based Atlases and a Database of Cortical Structure and Function*. New Jersey: John Wiley & Sons.

Wade, E., & Clark, H. H. (1993). Reproduction and demonstration in quotations. *Journal of Memory and Language, 32*(6), 805-819.

Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., & Ethofer, T. (2008). Cerebral processing of emotional prosody - influence of acoustic parameters and arousal. *Neuroimage, 39*(2), 885-893.

Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. In S. Anders, G. Ende, M. Junghoffer, J. Kissler & D. Wildgruber (Eds.), *Understanding Emotions* (Vol. 156, pp. 249-268). Amsterdam: Elsevier Science Bv.

Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., & Ackermann, H. (2005). Identification of emotional intonation evaluated by fMRI. *Neuroimage, 24*(4), 1233-1241.

Winkworth, A. L., Davis, P. J., Adams, R. D., & Ellis, E. (1995). Breathing patterns during spontaneous speech. *Journal of Speech and Hearing Research, 38*(1), 124-144.

Yao, B., Belin, P., & Scheepers, C. (under review). Brain 'talks over' boring quotes: Top-down activation of voice-selective areas while listening to monotonous direct speech quotations. *NeuroImage*.

Yao, B., Belin, P., & Scheepers, C. (2011). Silent reading of direct versus indirect speech activates voice-selective areas in the auditory cortex. *Journal of Cognitive Neuroscience, 23*(10), 3146-3152.

Yao, B., & Scheepers, C. (2011). Contextual modulation of reading rate for direct versus indirect speech quotations. *Cognition, 121*(3), 447-453.

Yoo, S. S., Lee, C. U., & Choi, B. G. (2001). Human brain mapping of auditory imagery: Event-related functional MRI study. *Neuroreport, 12*(14), 3045-3049.

Yuasa, M., Saito, K., & Mukawa, N. (2011). Brain activity associated with graphic emoticons. The effect of abstract faces in communication over a computer network. *Electrical Engineering in Japan, 177*(3), 36-45.

Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience, 14*(4), 1908-1919.

Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science, 256*(5058), 846-849.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. *Psychology of Learning and Motivation: Advances in Research and Theory, 44*, 35-62.

Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: dynamic representations in language comprehension. *Cognitive Science, 28*(4), 611-619.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168-171.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General, 135*(1), 1-11.