

**At the intersection of text and talk: on the
reproduction and transformation of language in
the multi-lingual evaluation of multi-lingual
texts***

WOLFF-MICHAEL ROTH

Abstract

Reported speech is one of, if not the most important feature in the reproduction and transformation of language as a *living* phenomenon – as opposed to F. de Saussure’s fixed semiotic system. Whereas there are quotation marks and clear grammatical rules in written language, speakers are known to draw on other, generally prosodic means to distinguish reported from reporting discourse. However, when the social situation involves the comparison of texts in two languages, a reporting context that permits both languages, and the co-presence of the original texts, then the question arises about the pragmatic solutions speakers ‘choose’ in a potentially confusing setting. The database for this study consists of think-aloud/interview protocols where 3 experts evaluate the official translations of 40 test items from a pan-Canadian science achievement test. A Bakhtinian (Vološinovian) approach is used to frame the levels of interpenetration of (indirect, direct) reported and reporting discourse, analyzed by means of prosody-including conversation analysis. Even in indirect speech, there is a prevalence to make reported words and expressions stand out by means of situationally adapted prosodic means; and, because of the co-presence of the source texts, extensive use is made of verbal and gestural deixis and iconic gesturing enhancing and even replacing direct quotation.

Keywords: indirect discourse; direct discourse; prosody; body movements; gestures; Bakhtin; Vološinov

1. Introduction

Hearing when a speaker reports the speech or text that someone else has produced is one of the mundane accomplishments of everyday interaction. Even when there is one predominant speaker, even linguistically untrained individuals can hear a speaker quote. This is quite evident from the transcriptions of think-aloud protocols with experts, who – in the same way as they do as part of their normal work – are in the process of determining whether the translation of a multi-lingual achievement test (completed by official, certified translators) is appropriate. A graduate student hired to transcribe the think-aloud protocol (but not trained in, or aware of, the topic of the present article) produces, among others, the following text.

It is again the the level of of language use, ‘lightning has struck’, very clear for an English speaking student. *La falaise a été foudroyée par un éclair*, how many student would be familiar with *foudroyée*.

Although the speaker has not used punctuation, we do note the commas and the period in the rough transcription. Even more pertinent to the present article, the graduate student has set apart two stretches of text, one by placing it between quotation marks and the other by italicizing it. The italicized text is easily recognizable as belonging to the French language, and the quotation marks suggest that the transcriber has heard to expert to been quoting. In the present, both set-apart texts are quotations or, in Bakhtinian terms (Vološinov 1930)¹, direct speech (*prjamaja reč*’). However, the stretches of direct discourse are embedded in evaluative comments, and, therefore, are part of indirect speech (*kosvennaja reč*’). Nevertheless, although speakers do not have these grammatical-typographical means to distinguish parts of sentence and quotations, the recipients of speech, like the transcriber, do hear clauses and quotations; and this hearing is captured in writing by the relevant means (italics, quotation marks, punctuation). What do

speakers do that allows recipients to hear that they are using direct speech? Moreover, what do speakers do to allow recipients to hear that they use indirect speech? What happens when the same textual features – words or expressions – recur repeatedly in a stretch of talk?

Reporting what someone else has said – e.g., for the purpose of supporting the current speaker's argument or as a topic to be commented on – is one of the most important phenomena of language, its tradition from generation to generation, and its continuous development (Vološinov 1930: 115): 'What we have in the forms of reported speech is precisely an objective document of this reception'. Although the sociological approach by Bakhtin/Vološinov 'did not inspire much systematic research into the interaction of grammar and style or the societal preconditions of different patterns of reported speech' (Coulmas 1986: 8), scholars interested in the micro-sociological (interactional) nature of societal phenomena (e.g., Collins 2004) should find the Bakhtin/Vološinov framework more suitable than other approaches. Scholarly texts, such as the present one, not only develop the discipline but also its discourse in part by means of directly quoting or paraphrasing previously published work. This leads to the heteroglossic (*raznorečie*, literally 'contradictory speech') nature of language, which develops precisely in the encounter of multiple 'voices', whether these come from different speakers or within the (outer, inner) speech of the same speaker (Bakhtin 1994 [1918]). Any appropriate understanding of dialogue therefore requires an understanding of the phenomenon of direct and reported (or quasi-in/direct) speech. Thus,

the productive study of dialogue presupposes a more profound investigation of the forms used in reported speech, since these forms reflect basic and constant tendencies in the active reception of other speakers' speech, and it is this reception, after all, that is fundamental also for dialogue. (Vološinov 1930: 115)

In written language, direct speech tends to be enclosed in quotation marks – though there are exceptions, such as in *Ulysses* (Joyce 1986). In spoken language, however, there are no literal

quotation marks, although one can at times observe, such as during conference talks, speakers who literally *read* a paper, mark the beginning and end of a quoted text by saying ‘[I] quote’ and ‘end quote’. Despite having been a phenomenon of scholarly interest for over a century (e.g., Vološinov 1930), direct and reported speech as features of face-to-face communication have gained prominence only more recently attention (Holt 2000) – in part because recording devices and analytic software makes the analysis more easily accessible. In living speech, a variety of prosodic means are used to mark that some part of what is being said is reported speech (Holt 1996, 2000). These markers, however, often do not coincide with the place where the quotation marks would be placed in written genres (Klewitz and Couper-Kuhlen 1999). Moreover, when the languages of reporting context and directly or indirectly reported text differ, an additional resource becomes available for demarcating the two forms of text, as shown in the present study. This additional language may, but does not have to, make it unnecessary to use prosodic means as an additional marker.

An important aspect of the ‘problem of reported speech’ is the evaluative aspect that is expressed in speech by intonation, that is, by prosodic means (Vološinov 1930). Just as Bakhtin/Vološinov describe, recent research now shows that speakers make available, through prosodic variations, their perspective and evaluation towards the reported text (Günthner 1999). Because the original text does not tend to be available, the degree to which a speaker imitates or modulates the prosodic features of the original speech cannot be verified. Thus, in most everyday instances, speakers cannot ground what they are saying in some public record. Only in special instances – such as in the case of parliamentary talk (e.g., Antaki and Leudar 2001) or recorded interrogations used in legal proceedings – can the veracity of the quotation be ascertained.

The complexity of the phenomenon is apparent in a cursory investigation of the following excerpt from a think-aloud/interview session with a translation expert, who, as part of his regular work, verifies and accepts/rejects translations of curriculum and testing materials that certified translators produced for a provincial government in Canada. Commenting on the problems in the equivalence of a question that had been part of a science achievement test administered by the

Council of Ministers of Education Canada – ‘Which is the most common cause of this type of damage to cliffs?’ and ‘Quelle est la cause probable de l’état de cette falaise?’ – the expert says:

Okay, type of damage, *l’état*, here ((pointing to French)), *de cette falaise* slight difference again. Here ((pointing to French)) we’re saying the state of the cliff, here ((pointing to English)) we’re already talking about what kind of state we’re talking about damage to the cliff right. So . . . so I haven’t read this ((covers with two arms the two versions of the test below)) yet, right, but here ((pointing to English)) I have a hint that that’s not normal, right. Okay. . . . Here ((pointing to French)) I don’t know *état de cette falaise*, I don’t know if its damaged or not. But here ((pointing to English)) we’re given a hint.

In these few lines, we observe a proliferation of ways of referring to a text, making it the object of the talk: there are direct reported text (‘type of damage’, ‘état de cette falaise’), indirect reported text (‘we’re talking about what kind of state’, ‘about damage to the cliff’), translation of direct reported text (‘state of the cliff’), gestural (pointing) and verbal deixis (‘here’), and complete uptake of a theme into the reporting speech (‘that’s not normal’, ‘I don’t know if its damaged’). Importantly, deictic gestures are used to make a piece of text the topic of talk even without verbally reproducing it. That is, in this brief excerpt, we observe an entire range of means by which recipients of a text take up the language of another, transforming and re-processing it until it has become part of their own repertoire. In fact, to understand ‘the specific nature of discourse as a topic of speech’, in the way it happens in this excerpt, requires ‘speak[ing] of another’s discourse . . . with the help of that alien discourse itself, although in the process, it is true, the speaker introduces into the other’s words his own intentions and highlights the context of those words in his own way’ (Bakhtin 1975: 167). Speaking of the discourse of another is possible only ‘when such discourse is utterly reified, a thing’ (ibid: 167), which is precisely what enables the gestural and verbal deixis as extreme form of quotation that we observe in this transcript.

In research on social phenomena, there always exists the question whether phenomena could be studied quantitatively by abstracting certain factors or features from the particulars of their settings to make them available for cross-situational study. Although there are studies that investigate, for example, the preparation in the reporting context of an anticipated quotation such as ‘be like’, ‘say’, ‘go’, or ‘think’ (Buchstaller and d’Arcy 2009), there have been suggestions that each (social, speech) situation is highly singular, never repeating itself (Bakhtin 1993). Something like a conversational ‘theme’ (*tema*), in contrast to signification (*značenie*), therefore is always unique (Vološinov 1930) so that investigations need to focus on the situated ways in which speech phenomena deal with problematic issues. It therefore does not surprise when some researchers state that even if quantitative counting could be done, ‘it would shed no light on how’ speakers use the words of others, ‘to what ends, and with what implications for their own and the other’s positions’ (Antaki and Leudar 2001: 469). If the phenomenon were a matter of making the direct reported speech stand out from and against the current state of the reporting speech, then local contingencies rather than global patterns would provide a sufficient account. For example, if the reported speech derives from a language other than the reporting speech, this may suffice as a distinguishing feature decreasing the necessity of using prosodic markers.

In this paper, I report on the phenomenon of direct and indirect reported speech in a pragmatic multi-lingual context where the original French and English texts are also available. In addition, as the experts involved also comment on alternative translations that would lead to a better equivalence between the two language versions and refer to what is more commonly said in one or the other setting, other forms of reported speech that Vološinov (1930) anticipated exist in the data for analysis, including that of not actually having to repeat the reported text but merely using verbal indices (e.g., ‘this’ or ‘that’), gestural deixis, or iconic and metaphoric gestures. The data derive from think-aloud protocols in which three experts (pseudonyms Paul, Sabina, and Walter), who, prior to their publication and use, verify and approve official English to French translations of official curriculum documents and provincial examinations made by certified translators. For two experts (Paul, Sabina), French is the mother tongue, whereas for the

third it is the second language (Arabic is the mother tongue). All three have worked for more than 2 decades in Anglophone (educational) settings but always with mandates for French language programming. Two experts predominantly thought aloud and spoke English (Paul and Walter), whereas one expert predominantly spoke French (Sabina).

In the present instance, the words and phrases used in reported discourse is held constant, for they existed in textual form without the performative, prosodic aspects of spoken discourse. This allows us to investigate the means in which speakers draw or do not draw on a variety of prosodic means when they report speech without ‘contamination’ from the original speech that is mimetically reproduced or transformed to impose social evaluation (Vološinov 1930).

The present study uses data from a very specific context: data collected as part of a project on the evaluation of translations of multi-lingual achievement tests. In this situation and in contrast to previous research on reported speech, the direct reported speech not only can be grounded in a (written) record but also is directly available to the participants, the translation experts and the researcher. Pragmatically, this provided additional means for bringing text authored by someone else into the reporting context (e.g., underlining, indexical and iconic gestures). At the same time, because the original text exists in written form, the inflections in the reporting speech are not mimetic reproductions unless they pertain to specific features of the original text – such as question marks or emphatic italics. Moreover, in contrast to previous research, the present data allow cross-language evaluations of the role of prosody, as the reported texts and the reporting contexts included both English and French. In this study, the methods of conversation analysis and interactional analysis of prosody (e.g., Couper-Kuhlen and Selting 1996; Goodwin et al. 2002; Roth and Tobin 2010) are used in addition to a computer tool for linguists that allows the determination of objective prosody parameters.²

2. Varying levels of interpenetration of direct and indirect discourse in multi-lingual reported and reporting contexts

The fundamental characteristic of language as a living phenomenon arises from the fact that it exists in use – a language is called dead precisely when no longer in use; and use implies a dynamic process, whereby language ‘exists asymmetrically, from the other, kept by the other, always returning to the other’ (Derrida 1996: 70). Thus, despite ‘having come from the other’, in every act of speaking it simultaneously will ‘have remained with the other’ (ibid: 70), to whom it will have returned. As part of this process, words, phrases, and longer stretches of discourse come to be attributed directly to someone else, who is quoted or paraphrased. Moreover, the speakers’ own voices modulate and accentuate not only the language deployed for their own intention – which may change for multi-lingual speakers in ‘certain “pragmatic” situations (rage or exclamation in familial and familiar situations, more often in private than in public . . .)’ (ibid: 77) – but also that which they report as having come from the other. In this process, there is a dynamic interaction between reported and reporting discourse: even in indirect discourse, where the original text is transformed for the purpose of the speaker’s intentions, ‘the dissolution of the reported utterance in the narrative context is not carried to its end’ and the ‘substance of the reported discourse remains detectable as a self-sufficient unit’ (Vološinov 1930: 116).³ Only in the never-achievable extreme would direct reported speech be inflexible, initially completely independent of the current speaker and, in some cultural-historical settings, remain that way. Real language – including written, oral, vocal, and musical, rhythmic, and prosodic gestures – is continually deformed and transformed, subject to expropriation, anomaly, and deregulation (Derrida 1996). In this section, I present different aspects of translation experts’ pragmatic solutions to the problem of reported discourse in the context of their evaluations of English-French translations of a Canadian achievement test.

2.1. Toward the study of communication: Toward integrative investigations of direct and indirect discourse

The research on direct and indirect reported speech tends to focus either on its implementation in texts (including Vološinov 1930) or on recorded *speech*. In everyday situations, however, communication is a much broader phenomenon mobilizing a variety of non-verbal means that together with the verbal means realize social interaction (e.g., Roth and Pozzer-Ardenghi 2006). When the text to be reported is physically present, speakers have the possibility to bring it into the reporting context – as a whole or in part – merely by pointing to, underlining, or gesturing along a piece of text. In the presence of the text as physical object, pointing, as iconic gesturing, constitutes an intelligible, signifying act because of

the mutual contextualization of a range of semiotic resources including at least (a) a body visibly performing an act of pointing; (b) talk that both elaborates and *is elaborated by* the act of pointing; (c) the properties of space that is the target of the point; (d) the orientation of relevant participants toward both each other and the space that is the locus of the point, and (e) the larger activity within which the pointing is embedded. (Goodwin 2003: 219)

Pointing achieves an interactive purpose when the recipient of the communicative act is co-present so that whatever is pointed to or iconically expressed is available. Especially relevant in the present instance is the mutual elaboration of pointing and talk: what is pointed to is itself a stretch of text that thereby comes to be quotatively integrated into the communicative act as a whole.

In the recordings analyzed here, because the experts and researcher are co-present, words entire phrases are imported into the communicative act – i.e., can be ‘quoted’ in an expanded sense – without actually being articulated. For example, gestures that begin over the actual French text and then move to the right far beyond in the course of talking about the extended length of this text or any other exam in which the French version would be longer than the English equivalent. Following the discussion of different ways in which the English and French

items present answers to the student – one only using a one-word answer, the other preceding the corresponding French word by ‘ce sont des’ (these are) – the expert uses pointing gestures while saying: ‘I’m a very practical reader and I like this ((pointing to the English item responses)) better than this ((pointing to the French item responses)) here ((pointing to the English version of the test)) where there’s less reading to do, there’s less reading to do than in French’. Here, the expert does not or, rather, no longer report the *text* but simply points to (parts of) and marks off the English and French versions by means of slashes, the former of which he likes better because it is shorter. In pointing, he thereby also points out not only the quotation ‘this’ but, because the pointing is to the English text, that it is the English rather than the French response form that he likes better (Figure 1b). This response has arisen from a preceding commentary two minutes earlier, where both *ergotic* gestures⁴ – i.e., drawing of vertical markers – and verbal reporting had been used. (Transcription conventions are found in the appendix; French text is to be sounded out in French.)

Fragment 1

```

01 R:      always talk out loud what youre thinking ok[ay: ]
02 P:      [yeap] [ ((marks
→          slash in front of 1st occurrence of 'ce sont des', Figure 1a))
03         pourquoi so=why:::: ah::: [ (0.29) why translate [ (0.99)
→          ((marks slash before [ 1st and after
          2nd occurrence of 'ce sont des'))
04         in [ snenglish here they give just the word [ right
→          [ ((points to beginning of 2nd response
          in English, Figure 1b))
→          [ ((points to end of
          4th response in English))
05         (1.77)
06         so=whY [add thIS [here ((00:10:54)
→          g [ ((Uses slashes to mark off 'ce sont des'
          in second of the four items, Figure 1a))
07         f <<acc,dim>[/ce sont dES/>
→          g [ ((slash before third item))
08         f <<all>/ce sont dES/[
→          g [ ((slash after third item))
09         /ce sont dES/.>> [ (0.40) [
→          g [ ((slashes before &
          after fourth item))

```

In this situation, there is no observable pause between the English reporting context and the reported (French) text; but there is a different solution to separating the two forms of text. The first instance of the reporting occurs by means of the ergotic gesture that marks off the second occurrence of ‘ce sont des’ in the text (Figure 1a), but without verbal equivalent. That is, the expert has already ‘quoted’ ‘ce sont des’ without actually having articulated it. The movement between the two instances of gestural deixis (turn 03) corresponds to the trajectory that the reading eyes follow, and it is therefore iconic to an aspect of the receptive process. While the explanation unfolds, he returns to the French part saying ‘why add thIS here’ and then, as if there were the need for repairing the indexical ‘this’, the expert then continues verbally reporting the text repeating the same identifying processes (turns 06–09) that already had marked the first occurrence of the phrase (turn 02). Thus, if there had been any doubt about what the first set of marks had accomplished, the repetition of the same movements concurrent with the verbal articulation of the (identical) phrases would have retrospectively solved any ambiguity.

«**Insert Figure 1 about here**»

Pointing and deictic terms tend to instruct recipients to attend to something other than the talk (Goodwin 2003). However, in the present, gestural and verbal deixis refers the recipient to a text that could have been said but which, here, exists in a different modality, in written form, perceptually accessible to speaker and hearer simultaneously. The deictic gestures, markings (underline, slashes), writings, and iconic gestures identify those textual fragments that in the non-presence of the written form would have to be imported into the communicative act in a different way. In the present situation, these gestures constitute a specific solution to the reception and uptake of language from another ‘in live dialogue’, where ‘we usually omit those words to which we are answering’ (Vološinov 1930: 115). In such situations, words and phrases generally are repeated ‘only in special and exceptional circumstances, when we want to check the correctness of our understanding, or trip our partner up with his words, or the like’ (ibid. 115). The present analysis shows that when additional opportunities for bringing the discourse of others into an author’s communication, such as pointing to a text, then speakers may draw on this

situation as a resource, 'importing' the text without actually articulating it by verbal means. As Fragment 1 shows, the relevant but non-articulated text may even be part of an evaluative discourse (turn 06), which exists here in the form of a question concerning the very presence of a phrase ('ce sont des').

The present example also exhibits yet another feature that may be used to distinguish reporting from reported speech: rhythm (Klewitz and Couper-Kuhlen 1999). In Fragment 1, there are no apparent differences in pitch and the speech intensity decreases progressively from turn 06 to turn 09. In the course of these four turns, the expert places marks at three of the four responses before the 'ce' and after the 'des'. However, there is a change in rhythm: from dam-da dam-da for the English reporting context (turn 06) to the da-da-dam, da-da-dam, da-da-dam of the reported discourse (turns 07–09). Together with the change in language, the result is that we hear the reported French text as distinctly different from the reporting (English) context. We have here a form of internal, rhythmic structuring that distinguishes the reported from the reporting speech similar to the case where both exist in the form of speech.

2.2. Cross-subject solutions to the problem of making noticeable direct and reported discourse

Throughout this article, 'semantic markers' are devices that speakers may use to set the direct and reported discourse apart from reporting discourse. The (changing) frequencies of such markers have been studied in different countries (Buchstaller and d'Arcy 2009) and the use of prosody to distinguish between direct and indirect speech in between- and within-speaker experimental designs (Nickerson and Chu-Charroll 1999). Other scholars suggest that phenomena such quoting and paraphrasing are subject to local contingencies (Antaki and Leudar 2010) and that every speech situation is utterly unique so that '*the most immediate social situation and the social milieu more broadly entirely determine – moreover, so to speak, from the inside – the structure of the utterance*' (Vološinov 1930: 88). This, then, would require completely locally relevant solutions to the problem of reported speech. Rather than grammar,

the informative nature of speech relative to the local situation appears to mediate emphatic accentuation (Bolinger 1972). The general absence of uniform strategies for differentiating direct and indirect discourse – and the marginal nature of the effect in a small number of cases – in within-speaker experimental designs (Kvavik 1986; Nickerson and Chu-Charrol 1999) may suggest that speakers resort to situated rather than general solutions to the problem. The present *naturalistic* study allows investigating any similarities and differences in which the problem of distinguishing reporting speech and reported text is solved because all three experts refer to the same texts.

In the present database, there are different ‘styles’ of speaking, where one expert (Paul) has a lot of run-in phrases whereas another expert frequently leaves marked pauses between words. This immediately suggests that the problem of making the reported text stand out is different for each speaker. The key moment, therefore, is ‘finding’ a situated way in making the reported text stand out against the reporting speech, which is not the same throughout so that the reported speech has to emerge from the local contingencies of current speech context. All speakers draw on the registers of possible ways of achieving the ‘goal’ of making the reported text hearably stand out. For purpose of exemplifying this issue, fragments from the three sessions are reported in which all three experts comment upon the same word or phrase. In the chosen example, the experts evaluate the first response option to the question about the factor that has the least effect on the loss of heat through the windows, which is stated as ‘Size of the windows’ and ‘la superficie des fenêtres’, respectively. All three experts comment on the non-equivalence of ‘size’ and ‘superficie’. All three provide a possible alternative to the English version: ‘la surface’, ‘la grandeur’, and ‘la taille’.

In Fragment 2, the expert (Paul) uses several different ‘techniques’ to make the reported discourse stand out from the reporting context, which is that of evaluating the equivalence between the two versions of the examination, that is, even though grammatically, he performs and achieves an indirect reported discourse.

Fragment 2

```

01 P: f <<f,dim>Size of the wINDows.> ((gazes at English text))
02 f <<dim>la superficIe des fenÊtres.> ((gazes at French))
03 (3.51)
04 [ i think=
→ 05 [f =↑sIze n sUperfic↑ie=
06 [ =is nOtdah (0.20) theresn (0.95) dIFference between
→ 07 [f sIze (0.26) superficIe, ((55.5 dB))
08 ((49.7 dB)) <<p,cres>right-> (1.17) <<p>and uh::> (1.25) all the
others are ^fINe. (1.79) <<p>so again theres; (1.61)
actually there is difference. (3.10)
09 u::h:a:> (0.67) its quite equivalent though
→ 10 [f but the ↑sI:ze= ((101 → 132 Hz)) =is::; ((25:27))
→ 11 [f when i read= ((55.3 dB →)) =sIze ((→ 63.3 dB)) i would
12 say the length and the width right
13 R: uh [hm ]
14 P: [not] the hEIght and the width. (1.04)
15 <<all>when=i=say>= ((66.0 max dB, 5.13 syl/s))
→ 16 [f =sUperficI:e= ((70.0 max dB, 3.76 syl/s))
17 <<pp,all>=its=a=bit=mOre itsa> ((53.0 dB max; 5.77 syl/s))
18 t sURface ((61.8 dB max; 3.53 syl/s))
19 ri[ght ] <<all>so i have to calculate> ((6.60 syl/s)) (0.66)
20 R: [uh hm ]
21 thIs: (0.31) tImes thIs whatever.

```

The fragment begins with apparent direct reported speech (reading). There are then four (attempted) comparisons (turns 4–7, 10, 11–14, & 15–21), where the equivalence of ‘size’ and ‘superficie’ is the theme. The direct reported discourse clearly stands apart, and is perceived as such given the bodily orientations toward the texts. In the first instance of a comparison, the two words at issue clearly stand out, as there is a substantial rise in pitch at the beginning of the figure (turn 05). On the other hand, when the same two words reappear (turn 07), they are in continuation of the preceding locutions; but there is a substantial drop in speech intensity (-5.8 dB) following ‘superficie’ as the expert comments on the equivalence of the remainder.⁵ There is a clearly noticeable emphasis on ‘size’ when it is articulated as part of an ‘aborted’ statement (turn 10); and the same word stands out again in turn 11, when the speech intensity increases more than six-fold (8.0 dB). There is something like a within-language translation when the expert suggests that reading ‘size’ corresponds to saying ‘length and the width’, which he corrects to ‘height and the width’. The French word at issue clearly stands out again, even though it is embedded in a continually unfolding, word-fusing utterance (turns 15–17) because of a

combination of prosodic changes. First, the speech intensity of ‘superficie’ more than doubles (4.0 dB) with respect to the preceding speech and is demarcated from the following speech by a 50-fold drop (-17.0 dB). There is also change in speech rate from 5.13 syl/s preceding the word, itself articulated at a rate of 3.76 syl/s, and followed by a rate of 5.77 syl/s. The translation of the word, ‘surface’, is equally made to stand out (turn 18): it’s articulation rate is 3.53 syl/s compared to the 5.77 and 6.60 syl/s of the preceding and succeeding locutions.

In the speech of the second expert (Walter), who is in the process of discussing the same two words, these come to be set apart prosodically as well as the proposed alternative ‘grandeur’ (turn 06) for ‘superficie’. Walter, who in other parts of the reporting speech has few accents, arrives at ‘la superficie’ (turn 04) having rhythmically accentuated the preceding speech now has to make stand out the reported word or its translation. First, the word ‘size’, articulated with falling and rising intonation and a bottom at the mean pitch before and after but a peak at 55 Hz higher than the surrounding talk. Furthermore, the reported ‘size’ is preceded and followed by the semantic markers ‘the word’ and ‘in English’, respectively. Similarly, we clearly hear ‘superficie’ apart, not only because it is a French term standing against the surrounding English, but also because of the substantial increases in pitch for both ‘la’ and ‘superficie’ and because of the semantic marker ‘en français’ that follows. The alternative proposed, ‘la grandeur’, equally stands out against the reporting speech: there is a more than four-fold increase in speech intensity (+6.3 dB) over the preceding and succeeding talk (turn 06).

Fragment 3

01	W:	and the difference that I have <<dim>noted is: ze difference between;> (0.29) dhe=word ((mean 135 Hz))
→ 02	f	↑~sIze ((max 190 Hz → mean 142 Hz))
03		↓in ENGLISH (0.42) which is (0.36) <<pp>sb> first (0.44) fairly straightforward and understOod by
→		/most=a (0.17) A:ge groups/ and
→		/most stUdents/
→		/most pEOple (.)/ whE:reAs=la ((2.88 syl/s)) (0.29)
04	f	↑la ↑sU:perficI:e: ((2.81 syl/s)) (0.14)
05		en français eha ↑mIGht be mOre of a=more difficult for some (0.25) <<dim>studENTS=e> ((last peak max 73.6 dB))
→ 06	t	<<f>la grA:ndEU:r> ((2 peak max 84.3 dB, avg 79.9 dB)) (0.28)
07		((77.8 dB, avg 73.2 dB)) probably would have been o mIGht

```

08 R: | have bEe:n a better term .hh (0.15) en in in in in
      | <<p>en> (0.11) my mind
      | <<h>so=its a> `small .h (0.17) negligible=[difference].
      | <<p>[uh hn ]>

```

In this case, there are in some instances jumps in the pitch level, particularly drawn out syllables, and within-word rhythmic accenting that does not exist in the surrounding speech. The reported words clearly can be heard as separated from the reporting speech, and yet in the three instances, this separation is achieved differently.

In the following fragment, the third expert (Sabina), like her peers, sets the reported speech apart prosodically, making it recognizable within the evaluative context typical of indirect discourse. In the first instance, her orientation and the movement of the index finger back and forth along the English text, then resting at its beginning, marks the locution as a direct reporting of what she finds on the page. It is embedded, however, in the commentary that she thinks the translation is correct, even though she notes that in the case of ‘size of the windows’, the text does not speak of ‘superficie’ (surface) but of ‘taille’ (Engl. size, height, waist). She contrasts this with the French – a contrast made apparent in the change in bodily orientation from one to the other language version – where the talk is about (‘ici on parle *de*’) ‘superficie’. The ‘size’, in terms of speech intensity, towers 21-fold over the preceding (+13.2 dB) French and more than 7-fold over the succeeding French speech (+8.8 dB). The entire word, with a u-formed pitch contour, is 30 Hz higher than the surrounding talk. The one-syllabic ‘taille’ (/taj/) is as extended as the 4-syllabic ‘on parle de la’ preceding it and as the ‘ici=on=parle=de=super[ficie]’ that follows it; we therefore hear it as *lento*, a marking that is further made audible by the step in pitch. The speech intensity of ‘superficie’ drops 5-fold with respect to the preceding talk, followed by a long pause (turn 10).

Fragment 4

```

01 R: <<pp>questce que tu regardes>
02 S: | <<p>no je pense que cest correct> parceque
→ 03 | f | <<f,h>size> of the win|dows|
      | | ((underline gestures along text | and back |, finger
      | | rests at beginning of English text))
04 | on parlepas là de la

```

→ 05 f ↑superficie ((*head moves toward French, then gazing at researcher*))
 06 on=parle=de=la
 → 07 t <<len>↑tAille> (0.30)
 08 ici on parle de ((*orients toward French version, points with pencil towards item*))
 → 09 f <<p>superficie> ((mean 64.7 → 57.4 dB))
 10 (0.94) mais je ne sais pas si ça fait une grande différence.
 ((mean 64.8 dB)) (2.17) parceque ((*turns gaze back toward English*))
 12 (2.43)
 13 R: <<pp>pourquoi> (1.08) <<p>dis=à=haute voix de>
 14 (0.34)
 15 S: <<f>bah>
 16 (1.65)
 17 R: comment on traduit ça d=une autre façon?
 18 (0.37)
 19 S: j=aurait peut-être ^dit uh:: la
 → 20 t ↑tAille ((170 → 270 → 173 Hz)) ((70.6 → 65.3 dB))
 21 ↓parce que la
 → 22 t ↑taille:e: ((*orients towards windows in the room*))
 23 <<dim>↓ça=peut=êt ça=peut=être des petites> <<cres>on parle pas
 vraiment> (1.77) ((*looks at windows*)) mais (1.93)
 <<dim>non=j=avoue que je sais pas.

In the second commentary (turns 19–23), the offered alternative to the official translation, ‘taille’, appears twice, each time prosodically enhanced against the ground, the reporting context. In the first instance, the mean pitch of 270 Hz compares to the 170 and 173 Hz that precedes and follows it, respectively; the speech intensity drops by more than a factor of 3 (-5.3 dB). In the second instance (turn 22), the word is twice as long as the first time, is raised in pitch above the surrounding ground, and is (pronounced 2-syllabically / taʃœ /) as long as the 6-syllabically uttered ‘ça peut être, ça peut être’ (/sa pø ʔɛt sa pø ʔɛt /).

In summary, all three speakers use prosodic means to mark both direct reported discourse, where an expression stands on its own (e.g., because it is read), and indirect reported discourse. The specific means, however, differ, both within and across the experts. That is, what appears to import is the change in parameter with respect to the context rather than a consistency across contexts. Thus, the word ‘size’ is articulated with a more than 7-fold intensity over ‘superficie’, the former being louder than the surrounding speech, whereas the latter is substantially lower than the preceding speech. Within the framework of BakhtinVološinov, the two versions of the

items are objectively available to all members of the societal setting. These constitute fixed authorial statements. The local problem of the translation experts is the comparison of the objective relation between the two language versions. The evaluation pertains to the work done by the translators, so that what the experts compare are the words and phrases *as stated*. That is, the experts do not compare alternative expressions that they are responsible for. Rather, they take the existing English and French expressions as stand-alone objects, the relation of which they evaluate. There is therefore little modification that the words and phrases permit, if the experts attempt to compare what their authors/translators have done rather than considering alternatives while attempting to produce a translation of their own.

2.3. Moving from direct to indirect discourse or The interpenetration of voices

Direct and indirect speech tend to be implemented differently, separating the former by means of pauses and higher initial pitch whereas the latter tends to take the pitch contour of a constative (Goodell 1987; Roth 2010). Vološinov, however, suggests that the boundaries are flexible; and because there might be a progressive uptake of another's discourse *in situ*, we can anticipate intonational changes associated with reported words and phrases. Because of the nature of the experts' task, there is a characteristic evaluative dimension that places direct discourse in the context of the evaluative reporting context. In a strong sense, therefore, most reported discourse – often unchanged and without the marker 'that' characteristic for indirect speech – occurs in the context of indirect discourse. An exception is the case of the experts' reading the texts aloud, which is a form of directly reporting what the texts say. For each of the 40 pairs of (English, French) test items that they have been asked to comment upon, therefore, the experts *first* report directly what the texts state before making evaluative statements in the course of which fragments – words and entire phrases – of the original texts reappear. Whereas in a written text there would be quotation marks or italics to mark that a word or phrase used is that from the texts discussed. To what extent do the speakers in the present setting mark such

Second, the direct reported (read) text ‘John reads about a characteristic’ begins with a 14-fold increase in speech intensity over the preceding talk; it is separated on both ends by clear pauses (0.65, 1.12 s). Third, the translation of the French item (‘Jean lit de nombreux livres traitants des caractéristiques des êtres vivants’) can be heard as distinct from the remainder of the speech because of the increased pitch level (Figure 2), a long pause following the text, a substantial change in speech rate (from 6.35 to 8.29 syl/s), and a 10-fold increase in speech intensity that follows the end of the phrase. As a translation, the phrase could be heard as both direct and indirect discourse, which is apparent from the fact that nothing would change if ‘that’ were to be inserted following ‘they say’. The expert here uses ‘they say’ to announce the coming of direct reported discourse. However, we already observe an interpenetration of direct and indirect discourse discussed below in the fact of the emphasis on ‘a lot’, which, because absent in the text (voice of its authors), derives from the voice of the expert. Such flags, however, frequently are missing, such as in turn 02; but it is apparent from other aspects of the setting – such as when an expert reads, writes, points to, underlines, or gestures along text – that we hear direct reported discourse. In turn 02, the expert follows the text with his index and small fingers, which allows us to perceive him as re-reading the text. These three cases illustrate the observation in the present database that *if* there are markers, then these fall together.

In a context where the reporting context is characterized by one language, a fragment appearing in another language clearly stands out even without prosodic and other marking. If a direct reported word or phrase is in the same language as the reporting context, then prosodic and other sign forms are necessary to make it stand out. In the present multi-lingual context, we may ask: Are there differences in the direct reporting of English/French text within a given (English/French) reporting language? Fragment 6 illustrates that in the present database we observe both absence and presence of considerable changes in one or more prosodic parameters alongside their absence in situations where the French reported text is embedded in an English reporting context. In the fragment, Walter is in the process of elaborating differences between the English and French version of a test item: ‘Which of the following would have the *greatest*

impact on lowering energy use?’ and ‘Quelle pratique entraînerait *la plus forte* réduction d’énergie utilisée?’ The first English direct reported text in turn 02 is clearly separated by pauses from the surrounding reporting speech, whereas the indirect reported text (turn 04) does not involve significant changes in the prosodic parameters. In the same way, the first French reported text (turn 06) does not involve a change, although following the second quotation, the reporting context drops to one-sixth of the preceding mean speech intensity (-8.1 dB, turn 07).

Fragment 6

```

01 W:      the dIFferENce that i=am=u (0.13) n (0.14) ~n:nOticing (0.16)
           i:::s=z:ee:::: (0.19) ^ENgl~ISh (0.79)
→ 02      f  ~grEAtest (1.19) IMPact (0.72)
03          bd (0.27) uh:m: (.) uh=uh its the ung=
→ 04      f  =grEAtest=t IMPact <<dim>on lOWering energy u::s:::e.> (0.55)
05          u:m: (.) its a cLEA:rer for again and thAT age group; (.) wa=um
           (0.37) 'then (0.20)
06      f  la plUs fORte (0.38) q=rÉductio:n (.) dénergie (0.48) utilisée:.
           f  (0.55) uh::: e:m: (0.13) <<cresc>plus fORte rÉducti:Oon>
→ 07      (-8.1 dB) (0.25) and the (0.21) sort of the contradIction
08      f  between (+5.9 dB) fORte and rÉductIO:n (+4.4 dB) m=mI:G:ht be
           confu:sing (0.34) uh uh for s=some uh (0.26) uh: u=u students
09          s so agAIIn (0.10) .hh (.) i I:: zink u::: the ENGLISH students
           a:r:e ~FAV`or=ed by:::: .hh ~cLEA:R`ly isolating (0.25)
→ 10      f  <<acc>greatest> (0.33) <<rall>'IM~pa:ct> (0.52) u:h u:h FROM (.)
→ 11      f  <<acc>sort of> (0.11) lowering Energy 'U:se (0.39)
12          <<rall>whe:r:ea:s:> (0.29) the frEnch translation jUMped (0.11)
           dIRectly: (0.19) .hh (0.19)
→ 13      t  (+ 8.6 dB)<<dim>the grEAtest redUction in energy Use;>
14          (+5.0 dB)↑~mI:ght lead to=s:::OME confUSion for s:OM:e (0.26)
           students.

```

We then observe a number of substantial changes in prosodic parameters that separate the English reported text ‘greatest impact’ (turn 10) and ‘lowering energy use’ (turn 11) from the reporting context, mainly involving substantial (two-fold) changes in speech rate. In turn 13, Walter delivers in English what has been announced in turn 13 as a rendering of the ‘French translation’. This translation of the French, now being in the same language as the reporting speech, is separated by pausing (0.49 s, turn 13) and a more than 7-fold (+ 8.6 dB) increase in speech intensity. This intensity drops toward the end of the phrase, and, again hearably changing, increases threefold (+ 5.0 dB) and by a substantial increase in pitch (+ 50 Hz) from the ‘use’ to

‘might’ (difference in mean pitch = 39.5 Hz). That is, whereas we hear clear changes in prosodic parameters when the reported text is in the same language as the reporting context but only occasional changes in the prosodic parameters when the reported text is in a different language.

In the present setting, the translation between language might sit between direct discourse, because what is said in another language is assumed to be the same, and indirect reported discourse, because the reported speech says differently than the original. There is a ‘hybridization’ (*gibridizacija*) (Bakhtin 1975) or intertextuality at work with continued shifting between the two original discourses interacting in the one of the reporting speaker. In such a situation, speakers get more easily away with fewer markers, as the change in language already marks the text. The difference is heard so that one might anticipate prosodic markers to be less necessary. In the present database, there are instances where pauses and changes in prosodic parameters make the reported discourse stand out as if it were direct reporting, although, as a whole, the grammatical structure that comes to be realized is that of the indirect discourse. The translation is not the same as the original, for the different ways in which some expression should be translated precisely is at issue in these sessions. The form of the reported text may actually contribute to this ambiguity. To see whether experts would use the same constructions, and, therefore, the same means for integrating and differentiating reported speech, we require specific forms of text that would require grammatical changes rather than simply becoming part of indirect speech by inserting ‘that’. For example, if the French were to state ‘j’étudie les caractéristiques des êtres vivants’, indirect speech would require a change in subject (and tense): ‘they say *that* he studies (studied) . . .’ However, there may not be a need to draw a strict boundary between direct and indirect reported discourse: There are suggestions that it cannot be made or is desirable, because ‘the narrative discourse advances upon the reported utterance, penetrating it with its own intonations’ (Vološinov 1930: 131) and because the reported utterance, in turn, ‘creeps into and disperses throughout the authorial context, rendering it flexible and ambiguous’ (ibid: 131).

2.3.2. *Additional boundary markers and the resolution of confounding situations arising from the multi-lingual nature of reported and reporting discourse*

Previous research has reported on the classical and newly arising forms of marking the beginning of reported speech, including, in English, ‘say’, ‘think’, ‘go’, and ‘be like’ (Buchstaller and d’Arcy 2008). These forms also exist in the present context, both to introduce direct reported text (‘be familiar with’, ‘translate’, ‘use’, ‘remember’, ‘know’, ‘on peut dire “la superficie des fenêtres”’ or ‘j’aurais peut-être traduit par . . .’, or ‘la traduction, la cause probable, c’est pas tout à fait pareil que “the most probable cause”’, ‘as translator, I would not have put . . .’, and ‘alors, si je comprends la traduction, “l’animal qui a laissé . . .”’) and indirect reported text (e.g., ‘is understood as’, ‘c’est juste dire qu’il existe des . . .’, ‘ça veut dire “à l’intérieur de la maison”’). In addition to the prosodic markers, all three speakers use semantic markers such as ‘the word’ (‘le mot’), ‘the concept of’, ‘the expression’, ‘l’expression’, ‘in English’ (‘en anglais’), or ‘in French’ (‘en français’). Sometimes these markers precede and sometimes they follow the reported term: ‘le même mot en anglais, skin type’, ‘en anglais il y a le mot: prothesis’, and ‘si on le dit en français’. As Figure 3 shows, other than the emphatically produced reported word ‘size’, there are no clear prosodic changes between the quoted word and its reporting context, unless the longer pauses between ‘between’ and ‘de word’, on the one hand, and between ‘in English’ and ‘which is’, on the other hand, are considered as markers of the kind that do not fall together with the place where written English normally would use quotation marks (Klewitz and Couper-Kuhlen 1999). Here, ‘the word’ and ‘in English’ are semantic markers so that the reported word ‘size’ is demarcated given the additional pauses before and after these markers. In the same way as quotation marks are collated to the first and last letter in English quotations (as in German but not in French), these semantic markers are ‘collated’ to the word that they are intended to distinguish. On the other hand, there are examples where non-separate word is repeated after a pause that then clearly separates the reported discourse from the reporting discourse (Figure 4). In the example, the definite article ‘la’ is initially produced

require distinct prosodic markers when they quote from the other language. However, repeatedly, a speaker continues in the language of the reported speech; that is, the reporting context now appears for a while in the language of the reported text. Thus, when the raw transcript reads ‘Why is it *élèves aux pluriel* and *participe* is not e n t [/*ø* *ɛ* n t e/]?’, we may ask how, if at all, the speaker uses prosodic means to make the difference between text and context hearable. In this situation, French is maintained as part of the main narrative in the same language as the reported discourse.

Fragment 7

01 W: [some students might take a few seconds to think why is it (0.14)
 → 02 [f élèves ((6.64 syl/s)) aux pluriel (0.13) and participe
 ((6.46 syl/s)) is not e n t. ((128 Hz)) (0.30) (138 Hz)
 03 nothing to do with the exam nothing to do with the question
 ((6.82 syl/s))

In this situation, because of the change from one to the other language, no other prosodic marker is necessary – theoretically; no other noticeable marker has been used in this situation, as the mean value of the parameters used here are not markedly different. We are in fact required to hear that only ‘élèves’ is the reported text, whereas its specification as a plural noun (‘au pluriel’) is not. However, if the economy rules in spoken discourse hold, then the word or phrase does not necessitate restatement: with respect to the content of the talk, there is a common history that includes speaker and recipient. Just as words do not have to be repeated in such situations (Vološinov 1930), markers do not have to be repeated to hear that a word belongs to one (e.g., reported) rather than another (e.g., reporting) discourse.

2.3.3. Sorting out concurrent within- and between-language translations

« Oui, je n’ai qu’une langue, or ce n’est pas la mienne. » (Derrida 1996 : 15)

The fact that I only have one language that never is mine, encapsulated in Derrida’s aphorism, lies at the very heart of language as a phenomenon, its constitutive role of

communication, its translatability within and across languages, its referential and signifying nature, (infinite) semiosis, and so on. Translation is a complex issue that occurs both within a language and across two languages (Benjamin 1972; Ricœur 2004). Within a language, a person translates from one expression to another following questions such as ‘What do you mean?’ The person queried then provides in different words what is taken to be the same. Such a translation corresponds to the intra-language translation (Derrida 1996).⁶ Translation constitutes one of the forms in which an original text is given further life: ‘In [translations], the life of the originals achieve their always renewed, latest, and most encompassing development’ (Benjamin 1972, p. 11). Translation is one of those processes that are fundamental to the life of language: ‘each element of the utterance that can be isolated and is endowed with sense and the entire utterance are translated into another context, active, answering context’ (Vološinov 1930: 104).

Within language, a translation may occur when a word or phrase is replaced by another one assumed to convey the same idea or information, such as when the recipient responds with ‘what do you mean?’ to an initial utterance by using different word(s) or phrases to say what is presumed to be the same content. In the present French/English setting, any potential ambiguity whether the reported speech is a within- or between-language translation may be resolved by additional markers. For example, the orientation may be consistent with one language version, which would make a reported speech more readily hearable as a within-language translation. On the other hand, if there is a bodily and gestural orientation from one to the other text – or even a metaphoric gesture that suggests translation from one to the other, then we (more likely) hear a between-language translation.

In the database, an expert sometimes apparently uses direct speech – as hearable in the prosodic changes – but the reported speech is actually a translation from the other language. That is, speakers and hearers have to resolve the problems of attribution of a stretch of reported speech to its source. This is so because an English reported speech may be referring to the original or to the translation from French, just as a French locution may be a direct rendering of a French original or a French translation of an English original. In the following fragment, we have

both types of translations occurring. Even though we can hear the differences between the direct and indirect reported speech and reporting context, additional resources are required to allow the distinction. Operating cross-modally, gestures that follow the written text constitute additional resources that allow us to hear the locution as something reported rather than reporting. In the case of reading, there is actually an orientation to the text rather distinct from the head and gaze orientation when the expert speaks to the researcher. The following fragment begins when the expert reads aloud the number of the current item and then the first sentence of the English and French versions. (Readers will note that the French direct report of the text is without the accentuation that the literature tends to suggest marks direct reports [e.g., Goodell 1987].)

Fragment 8

01 P: twentyfi:ve (0.48)
 02 f ↑`whAt ↑`kInda <<dim>↑`fOrce ↑`keeps the ↑`cnoe afloat.> (0.25)
 03 f <<dim>comment=s=appèle la force qui empêche les canots de
 couler.> (-11 dB)
 04 (+8.5 dB) <<dim>ha=ha=ha=ha> (amplitude 17 → 6.6 dB) (0.26)
 .Hhhhhh (0.73) ^hha::::: <<p>lookadhat>
 05 (1.33)
 06 R: 'why are you laughing;
 07 (0.64)
 08 P: ^s:o:=ah::::: (1.23)
 09 f 'whAT ↑`kinda ↑`force kEEps the cnoe 'afloAt. ((points to English
 item, follows with middle finger)) (1.39)
 → 10 t <<len>'whAT `kinda force prevENts the canoe from sIN:king.>
 ((follows the French item with his finger)) (0.58)
 11 R: <<p>uh hm>
 12 (0.47)
 13 P: okay (1.66) so again is different right?
 14 R: <<pp>hm> ((nods))
 15 P: <<p>okay> ((points to English)) (1.94) `so ^why ((moves pointing
 fingers to French)) not <<dim>u:s=e=a::::: > ((movement back
 to English and again towards French, as if "carrying" something
 across)) (1.53)
 16 [t so (+85 Hz) ↑`what dhe force that prevents the canoe from
 <<h>'sINK-ing.> (0.47)
 17 (-59 Hz) ↓you have to tink (.) yuh:. (0.91) its a brai::n:
 (0.70) its a different brai::n::: (.) <<dim>exerci:s:::=because
 its by the nEgative right.>
 18 t 'whAT prevENts from 'sINKing. (0.62)
 19 he:re::: ((points to English)) (2 syl/s)
 20 [f (-8dB) <<all>'whAT kEEps them afloat.> (4.37 syl/s)
 21 (1.27) ((researcher nods))
 22 <<p>so anyway.> (0.82) khkh (0.95)

```

23      [ i would=have=say:: ((points to English, then 'moves' hands
        to French)) (0.62) <<p>ye=know> (1.08) i would=ve say (0.58)
        ye=know ((hand movement from English to French))
        <<dim>i would say;>
→ 23      [ t (+34 Hz) <<dim,all>↑qu=Est=ce=qui permet aux canots de flotter.>
24          (0.44)
25 R:      uh=hm ((nods))
26 P:      <<pp>all=right> <<p>very simple.>

```

When asked, the expert then first points to the English text now, in contrast to the first reading, accentuating 'afloat' whereas two previously existing accentuations (kINda, fORce) have disappeared (turn 09). He continues to the French text and, while following it along with his middle finger, says in English 'whAT ↑`kInda ↑`fOrce keeps the canoe afloAt' (turn 10). We note that the intonation contours and speech intensities are very similar (though different from that of the reading in turn 02). This allows us to hear or anticipate the comparison of 'keeps . . . afloat' on the one hand and 'prevents . . . sinking' on the other hand, which is clearly seen when the two phrases are mapped against each other after adjusting for speech rate (Figure 5). There is a difference, though, in that the second, translated expression is produced at a slower rate, making it stand out against the former. There is also a difference in that we may (subjectively) hear the rising pitch in 'afloat' as ironical, whereas the rising falling intonation in 'sinking' can be heard as an affirmative. He then announces that there is a difference and launches into a comparative where the original texts are reported in indirect (original, translated) form (turns 16–20). We notice that the first production of the English translation accentuates only 'sinking' but when it is repeated, both terms in the double negative 'prevents' and 'sinking' are emphasized. That is, the emphases in the reported text mark the current topic, which is the double negative in the French version compared to the positive statement of the question in the English version. However, the second accent (i.e., on afloat) disappears when the English version of the figure is repeated (turn 20). We then see that he offers a French equivalent for the English version (turn 23). However, in this instance of indirect speech, all emphases but the one marking the question have disappeared. That is, prosody has been situationally employed to make stand out relevant aspects: the parallels or contrasts between English, French, and translated versions.

In fact, the phrase in turn 23 can be heard both as within-French translations, stating the same with different words, and across-language translation, here as a better rendering of the English ‘what keeps the canoes afloat?’.

««««« Insert Figure 5 about here »»»»»

The preceding example shows that translation can be made hearable through pitch contour repetition across the two languages so that there then exists both a boundary, between reporting English and the foreign reported French word, and an equivalence, as if the translation were only within language. In the following example, the ‘what is’, though it is the beginning of the next phrase on the English side, can be heard – at least after the fact – as a parallel to the ‘there are’ and ‘on trouve’, the prosodic contours can be heard to be the same (Figure 6). This is interesting in the context of the claim that only in a ‘loose/cowardly sense of the term “translation”’ does everything become translatable and that translation meets its limits in ‘prosody, metrics (accent and quantity in the time of pronunciation)’ (Derrida 1996: 102). In the pragmatic context of the task of comparing translations, the intonation across languages does in fact exhibit both difference and equivalence of the two languages. The alternative ‘[il] y a’ then appears as part of a question, which not only presents the alternative but also does so with an evaluative ‘why not say?’, a common instance of indirect discourse (Brown 1980).

Fragment 9

```

01 P:    <<pp>okay> (1.19)
02      t    <<rall>what <<dim>i:::s::: >> (0.16) ((gazes at English))
03          KH:: (1.04) [ u:::h <<all>yeknow>[ (9.05 syl/s)
04          [((mark after "are"))][((mark before "there"))
05          <<rall>there are> (2.05 syl/s) (0.27) <<rall>on [trouve>[ (2.48
06          syl/s) (1.30)
07          [((mark [ before
08          "on" and after "trouve"))
09          [why=not=say (.]
10          [((begins to write "il y a on top margin
11          y=a (0.19)]
12          of sheet))]
13          right? (-4.0 dB)
14          (0.30)
15 R:      u[h hm]
16 P:      [okay]

```


speech act that has already (actually) taken place. It may concern a hypothetical discourse or something that is or is not said in general culture or in specific contexts (family).

2.4.1. Subject-analytic modification in indirect discourse

We observe an example of the subject-analytic modification at work in the following excerpt from the analysis of a text, part of which is stated ‘The students travel by boat along the coast of New Brunswick’ and ‘Pendant une excursion en mer, les élèves longent la côte du Nouveau-Brunswick’. In this transcript, presented in the way the graduate student had transcribed it, the expert suggests that the translation is not word-for-word, that the English version explicitly states that the students travel by boat whereas the French version only notes that the students follow (‘longent’) the coast.

The students travel by boat. . . . Bon, ils sont en bateau, le long de la côte du Nouveau-Brunswick. Donc là ((French)), oui, « pendant une excursion en mer », on dit, ils longent la côte du Nouveau-Brunswick. Donc c’est, la traduction est un peu différente, mais ça veut dire la même chose, parce que c’est tourné de façon différente. « Pendant une excursion en mer », on dit pas qu’ils sont en bateau. Mais enfin, c’est forcément en bateau parce qu’on ne peut pas aller en mer autrement qu’en bateau. Donc, c’est pas tout à fait pareil, c’est pas traduit mot à mot, mais ça ne, je ne pense pas que ça pénaliserait. (Sabina, Item 6)

We note that the graduate student does not specifically mark the English phrase, which is a direct quotation, consistent with the analysis in the preceding section. The opening is followed by an assessment of what is being said, ‘ils sont en bateau’ (They are in the boat). The graduate student heard what follows as indirect discourse: ‘Donc là, oui, pendant une excursion en mer, on dit, ils longent la côte du Nouveau-Brunswick’. Although ‘pendant une excursion en mer’ and

‘longent la côte du Nouveau-Brunswick’ are directly from the item, they are part of a larger structure. There is a transformation from ‘les élèves’ to ‘ils’, preceded by ‘on dit’, which could have been followed by ‘qu’ (that), ‘on dit qu’ils . . .’ How are such instances achieved pragmatically so that a linguistically naïve recipient, such as the transcriber, will use the punctuation (comma, quotation marks)? What are the resources performatively supplied that lead to the kind of transcription provided?

In the following fragment, we observe the production of direct reported text to indirect reported text, with the special feature arising from the particularity of the setting that the presence of the text in physical form may be referred to by deictic gestures and the verbal indexical (shifter) ‘here’. The essence of this fragment is the comparison between the English and French version of an item, where the expert indicates that the former makes available information that is not available in its French translation, the supposed equivalent. As the second part of the query/reply turn sequence, the expert restates the texts that he is currently looking at, ‘type of damage’ (turn 03) and ‘l’état de cette falaise’ (turns 03, 05): these are instances of direct discourse clearly set apart by prosodic means. He then launches into an evaluation, which compares the translated French expression (‘state of the cliff’) with what the English text is ‘already talking about’: ‘damage to the cliff’ (turn 09).

Fragment 10

```

01 R:      <<p>whatyer thinking> (0.80) or looking at.=
02 P:      =↑`okAy
→ 03      f  ↑tYpe `of |'daMAg:e:| (1.30) |l=é:tAt (0.33)
              |                | ((places vertical marks before and after
              "type of damage," then turns to French))
              |((marks "l'état"
04          he:re: (0.11) ((65.3 dB mean))
→ 05      f  <<p,dim>de cette falai:se:::> ((52.7 dB mean)) ((gestures along
              the three words))
06          <<pp>okay> slight dIFFerence again.
              | (2.11) here we=re saying the ((10.0 syl/s)) ((65.5 dB))
              |((moves pencil tip above
07          t  ↑state of the ((3.98 syl/s)) (0.30) cLIFF:| (1.24)
              'l'état de cette falaise' holding it) |
→ 08          here ((points to 'damage' with rH pen and lH index finger))
              we=re already talking about (0.47) what kind of state (62.0 dB)
→ 09          <<p,dim>we:re talking about damage> <<pp>to the cliff

```

10 R: [((55 dB mean)) ri[ght>]
 <<p>[uh hm]>
 11 P: so:: (2.92) so i havent read this ((covers with two arms the two
 versions of the test below)) yet 'rIGht (1.12)
 → 12 [`but <<f>↑he::re> ((points to 'damage')) ((72.3 dB)) (0.20) i
 have a hint ((points to "damage")) (0.11) <<p>that:t:> (0.20)
 thats not normal rIGht ((gestures along 'many cracks')) (0.51)
 <<pp>okay> (0.33)
 → 13 [he:r:e ((points to 'il y a des nombreuses fissures')) i dont
 know (0.36) <<p>état de cette falaise> <<dim>right <<pp>here
 okay>> (1.80) <<p>i dont know if its damaged or not.> (0.34)
 → 14 [but here ((points to English, 'damage')) we=re given a hint.>
 ((58.3 → 56.7 dB, 7.69 → 5.63 syl/s))

The expert then elaborates the evaluation of the French and English versions (turn 12), relying on verbal and gestural deixis and iconic movement along a stretch of text to mark the items to be compared. As if the indexical marking were insufficient, we also observe, in turn 13, the appearance of the identical words from the source text – i.e., ‘état de cette falaise’ – set apart by means of the lower speech intensity. In the modulation of the speech intensity from normal to piano and pianissimo, the speaker mobilizes sufficient registers to allow a transcription as ‘here I don’t know, *état de cette falaise*, right. Here okay, I don’t know if it is damaged or not’. We can hear that the French is a quotation specifying and elaborating on the ‘here’ that precedes it. This, therefore, corresponds to the cases that Bakhtin/Vološinov discuss as indirect discourse with the original words enclosed in quotation marks. The case where the direct discourse appears within indirect discourse and is prepared by it ‘constitutes one of the many variants of indirect discourse in its picturesque treatment’ (Vološinov 1930: 129).

In this fragment, therefore, we observe a movement from direct to indirect discourse, from explicit articulation of an author’s words to the indirect discourse that comparatively evaluates multiple expressions, to the disappearance of the specific text because it is the recognized topic and no longer needs to be articulated. This is a complex performance in which we observe direct and indirect speech, translated versions of text from the other language, as well as gestural reference to a text that does not appear in the verbal mode. Speech is highly variegated, with a great range of speech intensities, modulations in pitch, and speech rate – to mention but a few of

the more easily accessible speech parameters. Because of this variegated nature, the setting apart of the passage from direct to indirect reported text is not mechanical, a point of substantial critique, but constitutes a passage of analysis, transformation, and reformulation; this movement is accompanied by the introduction of and crossing with appreciative and evaluative accents.

2.4.2. *Verbal-analytic modification in indirect discourse*

This type of modification through indirect discourse integrates the words and turns of phrase of others' discourse that 'characterize the subjective and stylistic configuration of the latter qua expression' (Vološinov 1930: 128). To bring out 'their specificity, their subjectivity, their typical characteristics', the reported words and turns of phrase are introduced in such a way that the relevant features clearly stand out to be noted. In the present database, there are repeated instances where stylistic configuration of the expression becomes the issue. In such cases, the experts suggest alternative words and phrases, often noting that the actual translation is more literary than everyday French, or where their personal preference would have been other.

So uh . . . I would not have translated this ((points to 'not' in English text)) in this way ((points to French equivalent)) here. I would've make it shorter, in here. Okay. I would have say, 'Laquelle des caractéristiques suivantes n'est pas' . . . uh see the italic here ((underlines 'not' in English, then boxes it)) is in ((boxes 'n'est')) two, right. ((boxes 'pas')) Okay. So I would try, I would not have translated this ((points to 'not')) like that ((points to entire French sentence)) here.

In this excerpt, the expert clearly comments upon a stylistic feature of the reported discourse, suggesting that he would not have translated it from English to French in the way that item developers/ translators had done. We might ask how speakers pragmatically implement this verbal-analytic modification in indirect discourse, and especially how prosodic and other

signifying features are employed to distinguish reporting and reported discourse. In such situations – directly arising from the task structure, which required them to compare and therefore evaluate pairs of corresponding texts – it is therefore not surprising to find the type of modification that Bakhtin and Vološinov discuss.

Fragment 11

01 P: f <<p,dim>'WHY do ALL cattail <<all>plants look similar.>>
 f <<pp,all,dim>Pourquoi les quenouilles se ressemblent=elles
 'tOU:tes.> (2.05)
 02 kH:::::m 'okAY:
 03 (0.78)
 04 R: ↑what <<dim>are you looking at now.>
 05 (2.51)
 06 P: [<<f>kh kh> (1.16) thats ((points to French phrase)) u::::::::::h
 (0.14) cOM:plex:::: sen:tence structur:::::e (1.65)
 <<dim>|rELatively complex way of saying;|>
 |((vertical mark before, then underline of 'toutes'))
 → 07 a (+11 dB peak) <<f>pour↑qOUI::::::::::> (+6 dB peak) (0.47)
 a ((writes 'toutes' above and right of 'pourquoi'))
 a ↑tOUT(0.26)es (1.04) les quenOUilles ressemblent=elles. (1.21)
 → 08 <<p>okay.> (1.22) <<dim>thats a really nice literary (0.26)
 ((moves body and both hands to right, in continuation of
 written phrase, indicating 'toutes' at end)) french structure.>
 (0.25) <<p>okay,> (0.96) you nEVER use that ((places index on
 'toutes')) in daytoday french.
 → 09 a <<all>you say> ↑pour<<dim>↑qOUI tOUTes les quenOUilles
 a ressemblent=elles.
 → 10 f <<p>right.> so WHY put (.) <<dim>↑tOUTtes (.) (-2.6 dB)
 thERe ((points to 'toutes')) at the end.>

At issue here is the expression 'pourquoi toutes les quenouilles se ressemblent-elles *toutes*', which the expert judges to have a 'complex sentence structure', constituting a 'relatively context way of saying' (turn 06). Immediately following the latter expression, he then articulates what the expression is stating but he does so in a different way. This different expression is supposed to be the same, is an alternative to, the one that the official translators had chosen, but it is different. We might gloss what is happening here in this way, which brings out the nature as indirect discourse: 'they are saying "X", but they are really saying "Y"'. Here, the hypothetical

or alternative ‘Pourquoi toutes les quenouilles ressemblent-elles?’ stands out in the same way as the words and phrases from the written text before the experts.

2.4.3. *Further transformations*

From the perspective that Bakhtin and Vološinov develop, we may anticipate a progressive incorporation of another’s words into speakers own voice. In the interaction of reported text and authorial context, the latter, corresponding to the voice of the speaker, ‘strives to disintegrate the compact and closed nature of the reported discourse, to resolve it, to obliterate its boundaries’ (Vološinov 1930: 119). In the process, the reported speech comes to be integrated into the voice of the speaker, becomes an integral part thereof – the very process by means of which language comes to be handed down and transformed and, thereby, is reborn and remains alive.

When the indirect discourse makes a particular word or expression the theme, subsequent uses may no longer require prosodic and other signs to mark that a term has had its origin in one of the two sources. The very fact that the word or expression has already figured during the reading-aloud part produced for the listening researcher might be taken as sufficient to identify the source all the while discussing its suitability as an equivalent to a French expression. In the following instance, the ‘artificial hand’ is a term from the English version of the text, but, as Figure 7 shows, it is not separated temporally from the surrounding reporting context. There is also continuity in pitch (F0) from ‘artificial hand’ to the reporting context that follows – though we note a slightly increased speech rate (4.37 syl/s) compared to the surrounding talk (2.78 syl/s). But here, the expert states that ‘artificial hand’ is ‘very clear to all students’, which can be heard as an assessment he might have made given his previously articulated experience as a teacher of science. That is, this assessment would then be relative to his experience rather than a direct comparison of the work of the translators. He then continues saying ‘*prothèse* might be [very clear] for most students’, which again is a statement that does not compare the equivalence of *prothèse* and ‘artificial hand’ but the degree to which it might be clear to the (Francophone)

those features are carried back from the written context to the speech context. In the texts that the experts were to evaluate, the authors had used emphases to highlight certain situations that were crucial for the selection of the answer, such as ‘all’ and ‘toutes’, ‘least’ and ‘le moins’, ‘greatest’ and ‘la plus forte’, ‘not’ and ‘n’est pas’ (n’est-elle pas’) or ‘ne . . . pas’, and ‘harmful’ and ‘nuirait’. An interesting question is whether the experts in reading and analyzing these texts would allow their recipient to hear that a particular word is emphasized or whether the emphasis disappears in the stream of words.

2.5.1. Reproduction of emphatic markers

Despite an early, seminal paper on the unpredictability of ‘accent’ (Bolinger 1972), there is a continued discussion concerning the presence, absence, and placement of emphatic markers – of interest especially to scholars working on text-to-speech problems (Hirschberg 2002). In the present study, because the experts first read the texts, it is possible to verify whether and how emphatic parts of the text are rendered in the actual speech. When a speaker uses direct discourse to render in the present situation what another speaker has said, the reproduction of prosodically realized emphasis constitutes a form of mimesis. The relation between the italicized word and the prosodically realized emphasis, however, constitutes a translation. In the following fragment, the expert reads from the text of item 36. We clearly hear the difference between the prosodic implementation of his comment (turn 01) and the reading of the English text. The locution begins with an emphasized interrogatory ‘why’, which is followed by a decrease in speech intensity and pitch to the end. However, the ‘all’ is emphasized in a similar way with an increase of the maximum speech intensity over the preceding and following word (+ 3.0 dB) and a pitch that is considerably higher than the surrounding words. In French, which is spoken with much lower speech intensity, the second part of the interrogative ‘pourquoi’ is emphasized, with a subsequent falling of pitch and speech intensity. The in the text emphasized ‘toutes’ is uttered with higher

peak intensity (+2.0 dB) and separated from the preceding speech by a micro-pause, so that we do in fact hear the emphasis.

Fragment 12

```

01 P:      <<p>yea thats pretty good> (0.85)
→ 02      f <<dim>↑^wHY do ALL cattail plants look similar.>
→          f <<p,dim> PourqOUI les quenouilles se ressemblent=elles (.
          tOUTes.>
03          (3.83) (+12 dB) <<f>UH:::>

```

Inherently, there is no difference between the prosodic implementation of the emphasis in the interrogative and that of the italicized term. In fact, here the questioning locution is implemented not by a rise in the pitch toward the end of the locution, corresponding to the question mark in the transcription convention but by the emphasis on the interrogative. The two emphases may therefore be heard as be of different kind, one realizing the expression as a question, the other to make a word (expression) stand out.

The reproduction of emphatic markers may actually constitute a personal style. In the present database, there are differences across the speakers and situations, for not in all cases does the emphasis become hearable. That the reproduction of emphatic markers may be a matter of style can be seen in the present instances, where words that are emphasized in the text are not consistently marked in the reading or reported speech.

Fragment 13

```

01 W: f <<all>jill=notices=that=there=are.>='mANY (0.62) cattail
      -`plANts around=the ↑`pONd (0.22) <<pp>.hmd> (0.28) they look
      'similar (+91 Hz) (0.62) to the cattail plants <<p>uh> (0.24)
→ that ^grO:ow (0.45) nEAR her schOOol. (0.60) ^wHY:=dO: a:ALl
      (0.38) cattail ↓_plAants (0.38) look (0.55) -↓sIMilar.

```

Here, the ‘all’ of interest is articulated with a mean intensity of 81.3 dB and an increase in pitch within the word of 21 Hz. The preceding words ‘why’ (76.2 dB, $\Delta = 77.5$ Hz) and ‘do’ (72.2 dB, $\Delta = 77.5$ Hz) also have strong and weak marking with respect to the ‘cattail’ that

follows (74.0 dB, $\Delta = 22.0$ Hz) or the ‘plants’ (76.1 dB, $\Delta = 34.0$ Hz). There are many other words that are much more accentuated than the ‘all’, which therefore does not stand out in an particular way. The resulting indistinction of ‘all’ is visible in the plot of pitch and speech intensity (Figure 8). In the French version, the ‘toutes’ disappears completely: ‘<<dim>se= ressemblent=<<h>elles>=tou(.)tes.>’. From the ‘ressemblent’ to the ‘toutes’ there is a decrease of -10.3 dB; and the pitch, which is 116 Hz at the end of the former, falls to 95 Hz for the latter word.

«*Insert Figure 8 about here*»

In the case of the third expert, Sabina, the ‘all’ and ‘toutes’ stand out again. We clearly hear the ‘toutes’, as there is an increase of 98 Hz in the pitch and an increase of 4.9 dB in speech intensity over the preceding words ‘ressemblent-elles’, itself in the context of a continually falling pitch level and a falling speech intensity: ‘<<dim>pourqOUI=les=quenOUIlles=se= ressemblent=elles.>=tOUTes?>’. In English, the ‘all’ is less pronounced: <<dim>whY: dO A:Ll cat (0.13) tail pla:nts look similar.>’.

2.5.2. *The authorial voice penetrates direct reported discourse*

Emphasis may be introduced in the re-reading of the text because the expert is in the process of highlighting a potential ambiguity that arises from the emphasized words, as discussed subsequently. The following fragment begins after the expert has re-read the English part of the opening of a test item: ‘Maria and Raphael collect a sample of frog eggs to study frog development’ and ‘Maria et Raphaël recueillent quelques œufs de grenouille afin d’en étudier le développement’. He then announces ‘another subtlety, another not subtlety, but another ambiguity’. He then first re-reads the English version and, in turn 01, now reads its French equivalent. As he re-reads the French version, he now emphatically stresses the ‘en’. As if this were insufficient, he again emphatically (and by circling it with his pen) articulates the adverbial

pronoun ‘en’ as complement of the verb to indicate an object before articulating what it ‘could mean’ ‘here’ (turn 04).

Fragment 16

01 P: f mariA et raphaël recueillent quelques oeufs de grenouille.
 (0.99) <<f>khkh> ((clears throat)) (1.34) <<dim>afin d=EN
 étudIER le développement.> (0.44)
 02 (+5.6 dB)so
 03 [EN] (.) ((circles the word with pen))
 04 hERE (0.26) ((points)) alright (0.75) [could mea:n:] ((draws
 arrow to “grenouilles”)) (0.24) <<all>the devlpmnt> (0.14) you
 want to study the development of the grenOUILLES
 <<dim>[OR of the] ((draws arrow to “œufs”)) EGGs.>
 05 (1.28)
 06 R: uhn okay.
 07 (0.78)
 08 P: you=re not sURe.

Here, the ‘en’ is emphasized repeatedly. This emphasis, however, is not distinct from that which appears in the next sentence of the text, where the expert emphatically articulates both ‘harmful’ and ‘nuirait’: ‘Which practice would be *harmful* to the developing frogs?’ and ‘Laquelle des pratiques suivantes *nuirait* au développement des grenouilles?’ However, there is a difference. In the fragment, the expert re-reads a phrase, and the adverbial pronoun comes to be emphasized only in this re-reading, having been indistinct part of the entire phrase in the first reading. Although the emphasis may be hearable in the initial reading, it may be confounded in subsequent talk about the marked word, when it becomes the topic of talk. Then the emphatic marker may be due to the fact that the word itself becomes the issue rather than its emphatic marker. This is apparent in the following fragment, where the expert critiques the construction of the French sentence, where he suggests a different placement of ‘toutes’. He first proposes an alternative French expression that would better correspond to the reading level of the targeted student population (turn 03). He explains that the existing translation has a ‘really nice literary French structure’, which ‘you “never” use in day-to-day French’. He repeats, now articulating the phrase more rapidly, the proposed alternative, and then follows with the question, ‘so why put “toutes” there at the end?’ (turns 07–09).

Fragment 14

```

01 I:      ↑`what <<dim>are you looking at now.>
02          (2.51)
03 P:      [ <<f>kh kh> (1.16) thats u::::::::::h (0.14) cOM:plEX:::: sen:tence
            | structur:::::e (1.65) <<dim>rELatively complex way of saying;>
04          | t (+11 dB peak) <<f>pour↑`qOUI::::::::::> (+6 dB peak) (0.47)
            | t ↑`tOUT(0.26)es (1.04) les quenOUilles ressemblent=elles. (1.21)
05          | <<p>okay.> (1.22) <<dim>thats a really nice literary (0.26)
            | french structure.> (0.25) <<p>okay,> (0.96) you nEVER use
            | that in daytoday french. <<all>you say>
06          | t ↑pour<<dim>↑`qOUI tOUTes les quenOUilles ressemblent=elles.
→ 07       | f <<p>right.> so WHY put (.)<<dim>↑`tOUTtes (.)(-2.6 dB) thERE
            | at the end.>

```

The ‘toutes’ of interest (turn 08) is weakly marked prosodically and by the micro-pauses that precede and follow the term. Moreover, given that its position in the phrase is the present theme, its previous emphatic articulations (turns 04, 06) may be attributable to marking it as the topic or to its emphasis. In fact, the former appears more likely. A while later then, the ‘toutes’ is more weakly marked because of its placement in the sentence structure rather than its emphatic employment in the original question.

Fragment 15

```

01 R:      <<dim>be cause of->
02          (0.59)
→ 03 P:      [ be:<<dim>CAU:se of the sENTence strUCTure putting; ↑`tOUTes
            | at the end hE:re-> (0.32) <<p>okay.>

```

This is an instance of between language translation, whereby an alternative solution is proposed to the official English-French translation. But it may also be a within-French translation, an alternative to the expression chosen, which is more literary and more complex than saying what the expert offers as a day-to-day French expression.

3. Discussion

Previous research has shown multiple prosodic parameters that may be deployed simultaneously for separating reported speech from the reporting context (Klewitz and Couper-Kuhlen 1999). In the present study, concerned with the phenomenon of (direct, indirect) reported speech in a multi-lingual context where the reported texts are (physically) available in the setting, a range of features not reported so far in the literature have been identified – most importantly including the role of bodily movements. In multi-lingual settings, shifts in language may suffice to mark what is reported text within the different language in the reporting context. The boundaries between reported and reporting speech are not fixed and marked, as some studies suggest, but any boundary that might occur tends to implement some situated need. Moreover, prosodic markers change to take into account the current need, which is different when a word or phrase is displayed again versus when it is displayed for the purpose of highlighting a problem. That is, the speakers in this study not only drew on multiple prosodic parameters but also on other parameters to make the reported text present within the ongoing communicative effort. The analyses show that there was not one consistent means employed across situations within one speaker or within the same situation but across speakers. Rather, the key requirement that speakers pragmatically implement, when needed, is to make the reported text stand out in situ adapting the means to whatever the current conditions.

The present multi-lingual context brings to relief the issue of within and across language translation, the former actually being of the same kind as the latter (Derrida 1996). Within-language translation are treated as forms of indirect speech, whereas between-language reporting of speech can be done by means of direct reporting of speech. The issue becomes more complex when within-translations are made in the language of the reported speech followed by translations into the language of the reporting speech.

The results of this study, which draws on empirical data in naturalistic context but with controlled variables, as these would be used in experimental designs, are in support of claims of the situated solution to the problem of accentuation (Antaki and Leudar 2001; Bolinger 1972;

Derrida 1996; Vološinov 1930). The study shows that within speaker, accentuation may be distributed and shift (a) across expressive modalities and (b) across time as the topic unfolds. Thus, an initial accentuation of an italicized term may disappear as the expert discusses features of the text and, in so doing, begins to emphasize a textual feature other than the initially highlighted one. To derive a general model of the use and shift of emphatic markers in direct and reported speech researchers may have to begin with naturalistic settings so that the effects of moving into controlled settings (speaking → reading, speaking in social relation, speaking for intentional communication) can be modeled from the perspective of the more comprehensive theory.

References

- Antaki, Charles & Ivan Leudar. 2001. Recruiting the record: Using opponents' exact words in parliamentary argumentation. *Text 21*. 467–488.
- Bakhtin, Mikhail M. 1981. *Voprosi literatury i estetiki*. Moscow, Russia: Xudošestvennaja literatura (English: *The dialogic imagination*, Austin, TX: University of Texas, 1981.)
- Bakhtin, Mikhail M. (1994). *Problemy poëtiki tvorčestvogo Dostoevskogo* [Problems of the poetics in the work of Dostoevsky]. Kiev, Russia: Next. (Parts reproduced in the English: *Problems of Dostoevsky's poetics*, Minneapolis, MI, University of Minneapolis Press, 1984) (First published in 1929)
- Bakhtin, Mikhail M. 1993. *Toward a philosophy of the act* (Vadim Liapunov, trans.). Austin, TX: University of Texas Press. (The original, *K filosofii postupka*, begun in 1919, is available at www.philosophy.ru/library/bahtin/post.html)
- Bakhtine, Mikhail [Voločinov, Valentin N.]. 1977. *Le marxisme et la philosophie du langage: essai d'application de la méthode sociologique en linguistique*. Paris: Les Éditions de Minuit.

- Benjamin, Walter. 1972. Die Aufgabe des Übersetzters. In *Gesammelte Schriften Bd. IV/1*, 9–21. Frankfurt/M, Germany: Suhrkamp.
- Bolinger, Dwight. 1972. Accent is predictable (if you're a mind-reader). *Language* 48. 633–644.
- Brown, Gretchen P. 1980. Characterizing indirect speech acts. *American Journal of Computational Linguistics* 6. 150–166.
- Buchstaller, Isabelle & Alexandra d'Arcey. 2009. Localized globalization: A multi-vocal, multivariate investigation of quotative *be like*. *Journal of Sociolinguistics* 13. 291–331.
- Collins, Randall 2004. *Interaction ritual chains*. Princeton, NJ: Princeton University Press.
- Couper-Kuhlen, Elizabeth & Margret Selting. Eds. 1996. *Prosody in conversation*. Cambridge, UK: Cambridge University Press.
- Derrida, Jacques. 1996. Le monolinguisme de l'autre ou la prothèse de l'origine. Paris, France: Galilée.
- Goodell, Elizabeth W. 1987. Integrating theory with practice: An alternative approach to reported speech in English. *TESOL Quarterly* 21. 305–325.
- Goodwin, Charles. 2003. Pointing as situated practice. In S. Kita (ed.), *Pointing: Where language, culture and cognition meet*, 217–241. Mahwah, NJ: Lawrence Erlbaum Associates.
- Goodwin, Charles, Marjorie H. Goodwin & Malcah Yaeger-Dror. (2002). Multi-modality in girls' game disputes. *Journal of Pragmatics* 34. 1621–1649.
- Coulmas, Florian. 1986. Reported speech: Some general issues. In Coulmas, F. (Ed.), *Direct and indirect speech*, 1–28. Berlin, Germany: Mouton de Gruyter.
- Günthner, Susanne. 1999. Polyphony and the 'layering of voices' in reported dialogues: An analysis of the use of prosodic devices in everyday reported speech. *Journal of Pragmatics* 31. 685–708.
- Hirschberg, Julia. 2002. Communication and prosody: Functional aspects of prosody. *Speech Communication* 36. 31–43.

- Jahandarie, Khosrow. 1999. *Spoken and written discourse: A multi-disciplinary perspective*. Stamford, CT: Ablex.
- Joyce, James. 1986. *Ulysses*. New York, NY: Vintage.
- Klewitz, Gabriele & Elizabeth Couper-Kuhlen. 1999. Quote – unquote? The role of prosody in the contextualization of reported speech sequences. *Pragmatics* 9. 459–485.
- Kvavik, Karen H. 1986. Characteristics of direct and reported speech prosody: Evidence from Spanish. In Florian Coulmas (ed.), *Direct and indirect speech*, 333–360. Berlin, Germany: Mouton de Gruyter.
- Luciani, Annie, Jean-Loup Florens, Damien Courousé & Julien Caster. (2009). Ergotic sounds: A new way to improve playability, believability and presence of virtual musical instruments. *Journal of New Music Research* 38. 309–323.
- Nickerson, Jill S. & Jennifer Chu-Charroll. 1999. Acoustic-prosodic disambiguation of direct and indirect speech acts. In *Proceedings of the XIV International Congress of Phonetic Sciences*, 1309–1312. San Francisco, CA.
- Roth, Wolff-Michael. 2010. *Language, learning, context: Talking the talk*. London, UK: Routledge.
- Roth, Wolff-Michael & Lilian Pozzer-Ardenghi. 2006. Tracking situated, distributed, and embodied communication in real time. In Michael A. Vanchevsky (ed.), *Focus on cognitive psychology research*, 237–261. Hauppauge, NY: Nova Science.
- Roth, Wolff-Michael & Ken Tobin. 2010. Solidarity and conflict: Aligned and misaligned prosody as a transactional resource in intra- and intercultural communication involving power differences. *Cultural Studies of Science Education* 5. 805–847.
- Sacks, Harvey. 1974. On the analyzability of stories by children. In Roy Turner (ed.), *Ethnomethodology: Selected readings*, 216–232. Harmondsworth, UK: Penguin.
- Selting, Margret, Peter Auer, Birgit Barden, Jörg Bergmann, Elizabeth Couper-Kuhlen, Susanne Günthner, Christoph Meier, Uta Quasthoff, Peter Schlobinski & Susanne Uhmman. 1998. *Gesprächsanalytisches Transkriptionssystem*. *Linguistische Berichte* 173. 91–122.

- Todorov, Tzvetan 1984. *Mikhail Bakhtin: The dialogical principle*. Minneapolis, MN: University of Minnesota Press.
- Vološinov, Valentin N. 1930. Marksizm i filosofija jazyka: osnovnye problemy sociologičeskogo metoda v nauke o jazyke. Leningrad, USSR: Priboj.
- Voloshinov, Valentin N. 1973. *Marxism and the philosophy of language* (L. Mtejka & I. R. Titunik, Trans.). Cambridge, MA: Harvard University Press.

Notes

* The data collection was supported by a grant from the Social Sciences and Humanities Research Council of Canada. The Griffith Institute for Educational Research provided support for proof-reading.

¹ The source of this text has been subject to considerable debate. It was originally published under the name of Vološinov (1930). The French version (Bakhtine [Volochinov] 1977) used here together with the original has been attributed to Bakhtin, referring to the contribution of Vološinov only parenthetically. Some linguist insiders, such as R. Jakobson (preface to Vološinov 1930) or Todorov (1984), believe not only that Bakhtin produced much of the text but also that the French translation is better than the English version (Voloshinov 1973). Here, I draw on and reference the French translation and the original. Here I cite and quote from the French version with page numbers of the Russian original in square brackets.

² The software, PRAAT, has been prepared for a variety of platforms and can be downloaded for free from www.praat.org.

³ An utterance (Rus. *vyskazyvanie*) is a text that is not limited to a speaking situation: A novel or poem constitutes an author's utterance, which, as a whole, enters the process of social evaluation (Bakhtin 1975). The utterance sits between and yokes together an author – text, speech act – and an addressee/recipient to whom, qua member of a social group, it is geared

(Vološinov 1930). The identifying characteristic of the utterance is the nature of a text in social situation.

⁴ An ergotic gesture is a movement that gets something done and simultaneously serves semiotic purposes (e.g., Luciani et al. 2009).

⁵ Each 3 dB constitute a doubling, each -3 dB a halving of the speech intensity.

⁶ Derrida (1996: 25) writes: ‘Car cette double postulation,

– *On ne parle jamais qu’une seule langue...*

(*oui mais*)

– *On ne parle jamais une seule langue...,*

ce n’est pas seulement la loi meme de ce qu’on appelle la traduction. Ce serai la loi elle-même comme traduction.’

⁷ Vološinov (1930) uses the terms *predmetno-analitičeskoj* and *slovesno-analitičeskoj*. The Russian *predmet* corresponds to the German ‘Gegenstand’, object, distinct from ‘Objekt’ (Rus. *ob’ekt*), object. The French *objecto-analytique* and *verbalo-analytique* correspond more closely to the Russian equivalents than the more loosely translated *referent-analyzing* and *texture-analyzing* of the English version. I use the adjective ‘subject-analytic’, because the analysis focuses on the semantic content of the utterance.

Appendix: transcription conventions

The transcription conventions are those of standard conversation analysis enhanced for the transcription of prosodic features (Selting et al. 1998). Unless modified, all words are written with small letters.

Notation	Description	Example
(0.14)	Time without talk, in seconds	okay. (0.24) OH
(.)	Period in parentheses marks a hearable pause less than 0.1 seconds long	NO. (.) <<assertive>you
(6.35 syl/s)	Speech rate measure in syllables per second	things.> (6.35 syl/s)
([+, -] 6.5 dB)	(Increase, decrease in) speech intensity, measured in dB	(+9.6 dB) <<h>here
((turns))	Verbs and descriptions in double parentheses are transcriber's comments	((head sidward))
::	Colons indicate lengthening of phoneme, about 1/10 of a second per colon	U::M:::
[]	Square brackets in consecutive lines indicate overlap	J: wa[xes.] C: [wAX?] (0.93)
[<u>OR of the</u>] ((draws arrow))	Underlined text in square brackets followed by description of a movement indicates extent of the movement described	[<u>OR of the</u>] [((draws arrow))
<p> >	Piano, lower than normal speech volume	<<p>like we get hot air from.>
<<pp> >	Pianissimo, a lot lower than normal speech volume, almost inaudible	<<pp>kay.>
<<all> >	Allegro, words are uttered with faster than normal speed	<<all>okay?>
<<len> >	Lento, slower than normal speech	<<len>it has (0.22) bin=decided that it is too dIFficult.>
<<plaintive> >	Transcriber's glosses are provided for	<<plaintive>i=just=found=

	ways of speaking	out=a=way=to='dO=`it.>
JAnE	Capital letters indicate emphasized sounds.	change of plA:Ns
.hh	Noticeable in-breath	..HHH UUm (0.42)
hh	Noticeable out-breadth	
-,?;. .	Punctuation is used to mark movement of pitch (intonation) toward end of utterance, flat, slightly and strongly upward, and slightly and strongly downward, respectively	T: so can we tell a shape by its color? T: does it belong to another group (0.67) O:r.
↓↑	Downward and upward arrows indicate pitch jump downward and upward	(0.36) ↓'O::N::; ↑IT (0.30)
=	Equal sign indicates that the phonemes of different words are not clearly separated	loo::ks=similar
˘	Diacritic indicates movement of pitch within the word that follows – down-up, up, down	˘buthh
→, f, t, [To the left of the transcription, arrows highlight specific turns; f marks direct figure of speech reported; t marks translation; [marks indirect discourse	→ 01 P: [f <<dim>'wHY did they sa:y::: des nombreux ^lI:::Vres.>

Captions

Figure 1. Two examples of the integration of reported text into the reporting context.

(The ‘outliers’ of the pitch are due to the letter ‘s’ in ‘say’ and ‘books’.

Figure 2. Gestures, underlining, writing, and other movements toward the physical text constitute important semiotic resources in establishing the equivalence of translations.

Figure 3. ‘The word’ and ‘in English’ may be counted as semantic markers, enhancing the function of the pauses to separate reported text (‘size’) from the reporting context (‘the difference between . . . which is’) in which it is embedded.

Figure 4. ‘En français’, which belongs to the reporting context, is produced in the same language as the reported text. The inhalation, pause, and interjection separate the semantic marker from the surrounding reporting text.

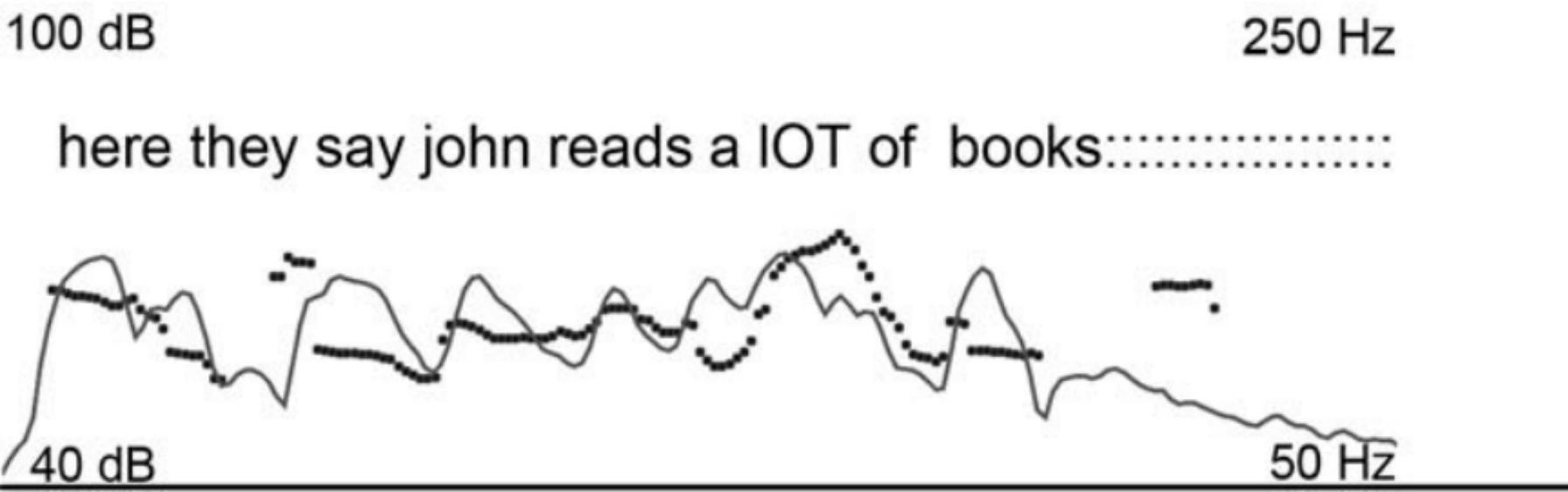
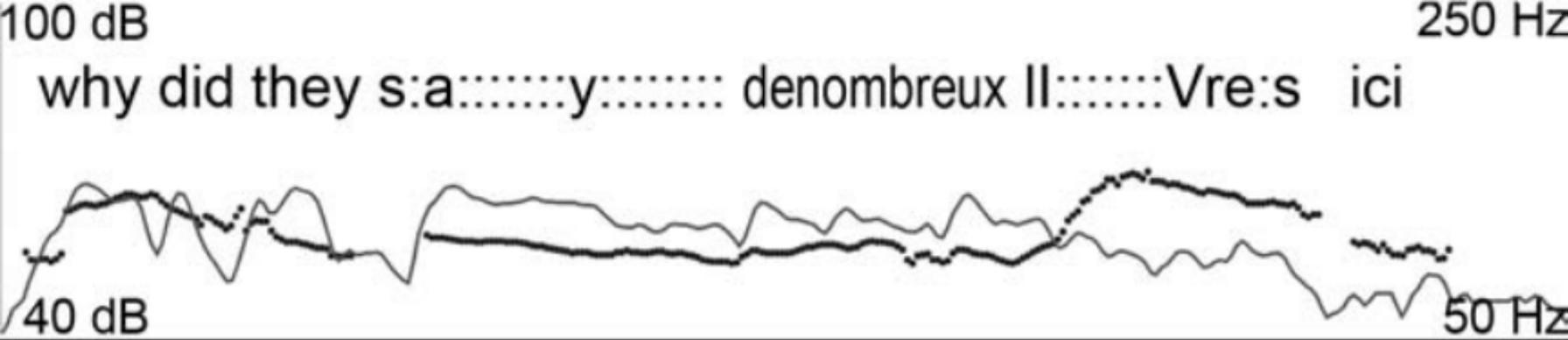
Figure 5. This comparison of an English phrase and the translation of its French equivalent shows how the current topic, the comparison between ‘keeps . . . afloat’ and ‘prevents . . . from sinking’ is made audible. (The lower phrase, which had been uttered *lento*, has been condensed to match the duration of the preceding phrase.)

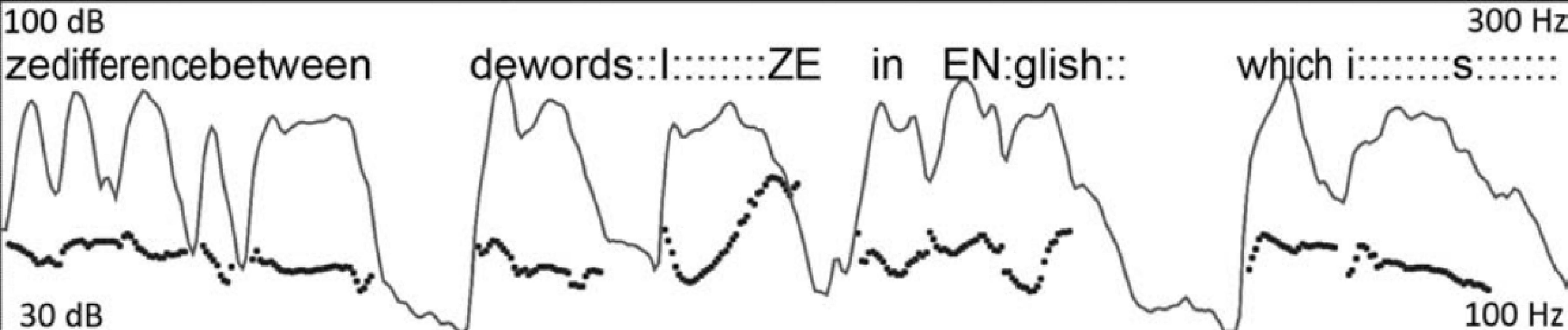
Fig 6. As the experts get into the discussion of a word or expression (‘artificial hand’), it, after originally being specially marked as reported discourse, increasingly is integrated into the reporting context. It has been completely absorbed into the reporting voice.

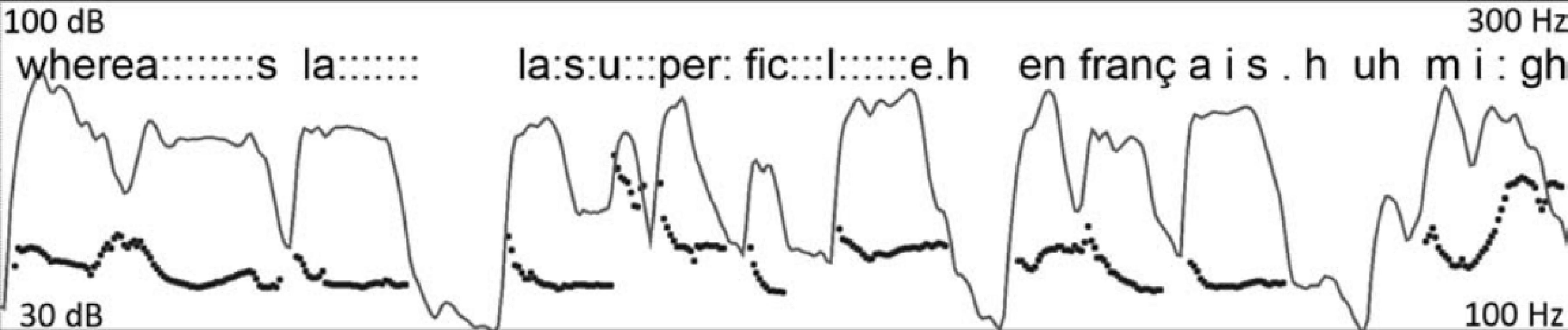
Figure 7. The equivalence of translations becomes hearable in the temporality (rhythm).

Figure 8. ‘All’, which is emphasized in the text that the expert reads, does not stand out any more than other words in the surrounding speech. It therefore is not heard in the emphatic way that the authors of the test item designed it to be.





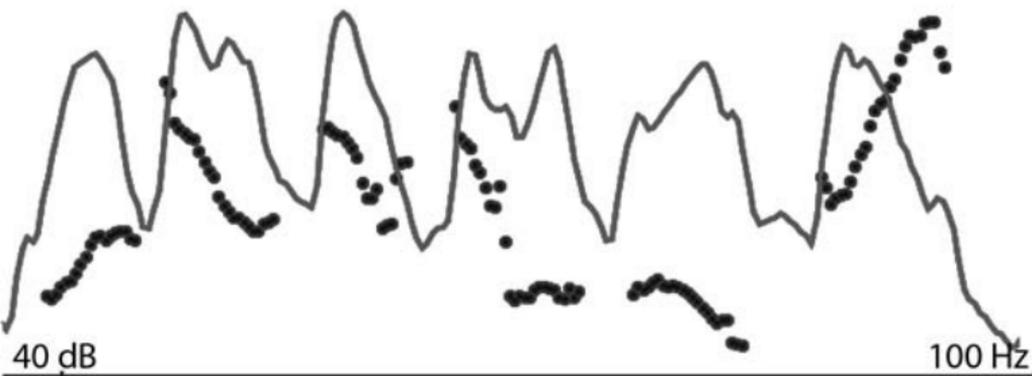




100 dB

300 Hz

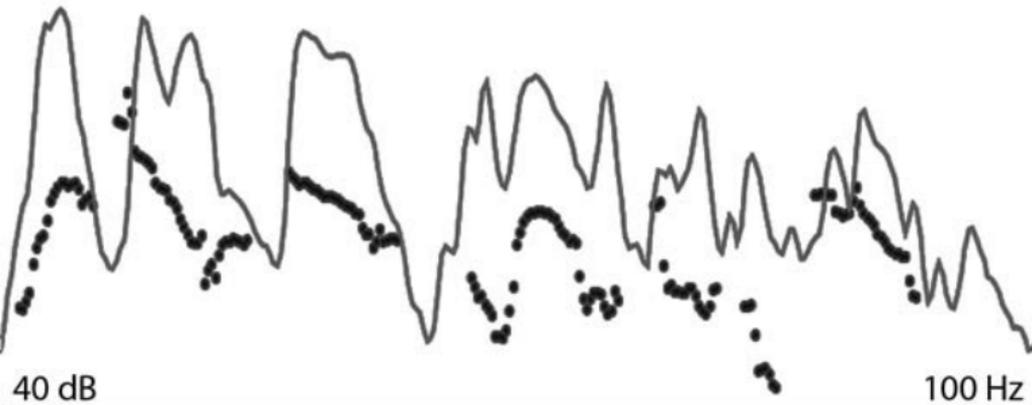
what kinda force kEeps the canoe afloat



100 dB

300 Hz

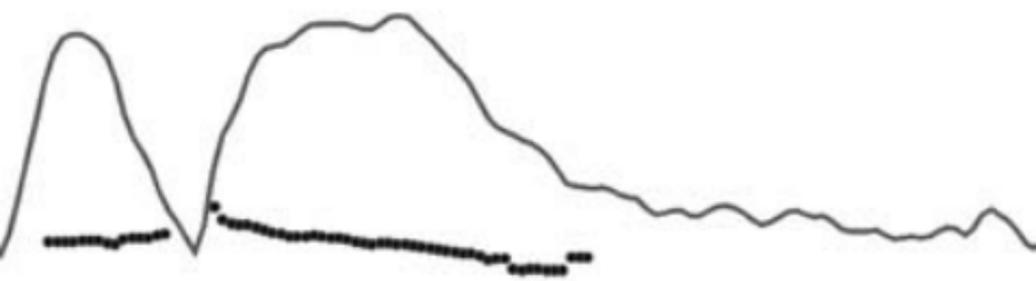
what kinda force prevENTsthe canoe from INking



40 dB

100 Hz

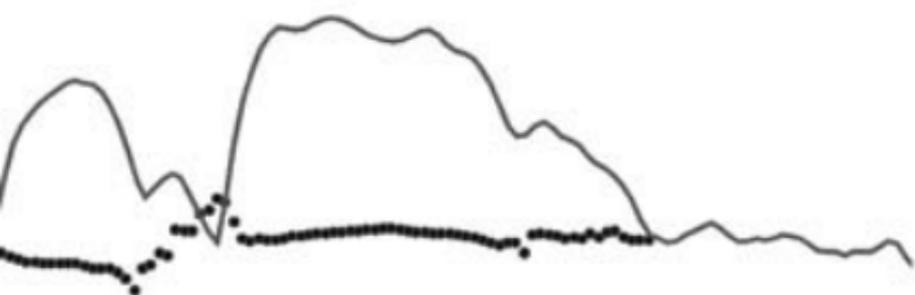
what i s s



there a r e



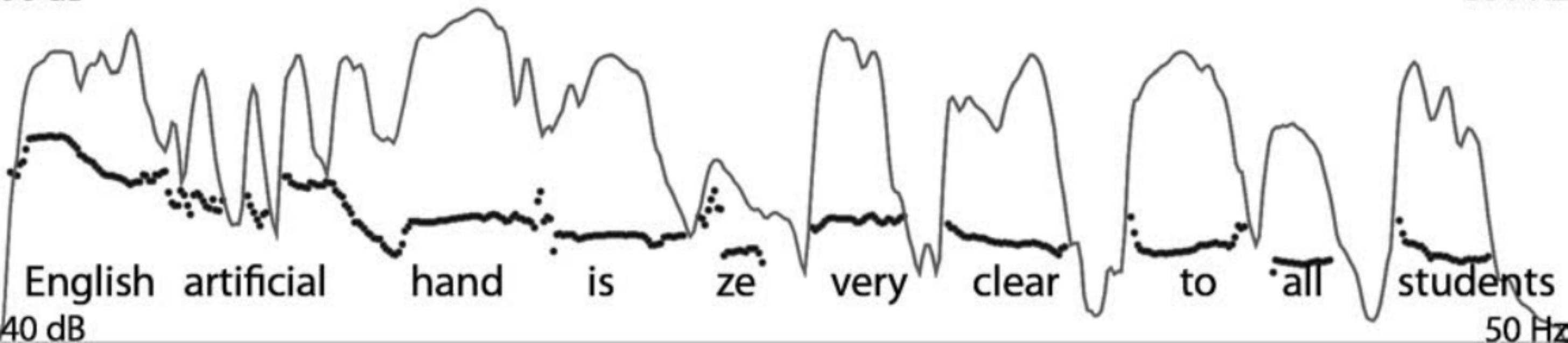
on tro u v e





90 dB

350 Hz



40 dB

50 Hz

