Contribution of prosodic and paralinguistic cues to the translation of evidentiary audio recordings

Raymond Chakhachiro
Western Sydney University
r.chakhachiro@westernsydney.edu.au

Abstract: This study examines accuracy in the translation and transcription of evidentiary audio recordings in the Australian context. Verbatim translation requested by crime agencies and courts is investigated and translation and transcription methods are suggested with reference to conversation analysis. The purpose of evidentiary audio recordings dictates a faithful translation; however, the prevalent 'written to be read' translation and transcription styles used by crime agencies can jeopardise the output, given the problems created in reflecting the speakers' intentions, moods, power and attitudes. The credibility of transcripts when tendered in evidence in court hinges on the quality of the translation. In addition to the stylistic accuracy of the translation of speakers' interactions, the present paper argues that important discursive information exhibited in the suprasegmental features in conversation should be documented on transcripts, including prosodic and paralinguistic elements, such as intonation, timing of responses and volume. When strategically used, these features can help in placing the last pieces of the jigsaw puzzle, and producing 'audible', 'written to be read as if spoken' texts.

Keywords: Evidentiary audio recordings, transcription and translation, conversation analysis

1. Introduction

Translation of audio recordings is the transfer of meaning from one spoken language to another, which involves the transcription or conversion of speech into a written text. It is a unique field and mode of translation, which is concerned with speech-types having their specific purposes that dictate the translation and transcription approach to be adopted. Audio materials requiring translation and/or transcription include monolingual recordings for research purposes (e.g., Gumperz, 1996; Sacks, 1995), recorded statements (Edwards, 1995; Teichman, 2000) and bilingual material for contrastive analyses (e.g., Bolinger, 1989).

Evidentiary audio recording (EAR) is the spoken material recorded by crime agencies, using various methods, such as listening devices and telephone interception, to track suspected criminal activity. This material may be used as evidence in criminal or civil litigation. Where EAR material is spoken fully or partially in a foreign language, e.g., Arabic in Australia, its transcription involves translation into the official or default language of the

country where the material is to be used – English in this context. The transcription and translation (translation for short) of conversation spoken in another language is an established, specialised sub-field of court 'interpreting' (Edwards, 1995). However, despite the potential implications of translations tendered in evidence, and the likelihood of their being scrutinised by the defence and rigorously contested in cross-examination, the scarce literature on the translation of EAR hovers around technicalities and presentation (e.g., Edwards, 1995) rather than the pertinent notion of accuracy. This study attempts to address this deficiency and contribute to a further understanding of the meaning of accuracy in the translation of evidentiary recordings.

1.1 Deficiencies of the prevalent translation practice

In the prevalent translation practice, conversation is treated as written text. Grammar, syntax, structure and vocabulary in conversations in languages other than English (LOTEs) are usually standardised, which obscures or distorts the interlocutors' intentions and sociological information embedded in style and register.

Having the foreign language transcribed alongside its English translation (cf. Edwards, 1995), without having systematic recourse to and documentation of the conversation's prosodic and paralinguistic cues, is marginally useful when it comes to refreshing the memory of the translator or the bilingual reviser of the original in cross-examination. However, this exercise runs the risk of treating the original as a written text in the translation process and in court.

The predominant verbatim translation or, at best, 'semantic translation' (Newmark, 1988), devoid of proper encoding of prosody, may provide meaningless or ambiguous output when it lends itself to more than one interpretation and thus may give rise to endless arguments in court among litigants.

A functional-pragmatic translation approach (House, 2001) or communicative translation method (Newmark, 1988) – subject to clarity of the recording, and availability of context and co-text – take into consideration the prosodic features in the listening process but without having them actually documented. This could attract legitimate argument by the defence and often be deemed inadmissible in court on grounds of subjective interpretation. What is more, the principle of accuracy of court interpreting based on pragmatic considerations (Berk-Seligson, 2002; Hale, 1996) does not assist the translator of EAR, who, in addition to the potential lack of access to context and background knowledge, has no recourse to feedback, repetition, clarification, and, most importantly, the kinetic cues of conversation.

Unlike interpreting, translation and transcription proper, and despite the demand for and importance of EAR translation, the field is acutely under-researched and lacks technical and ethical standards. A freelance translator can be forced to follow totally different sets of in-house translation guidelines depending on the agency. In the absence of commonly agreed-upon guidelines, translators for the same agency can have different understandings of the translation method required due to the lack of systematic training. This situation becomes more serious when the defence engages an independent reviser as expert witness.

EAR materials can be riddled with exophoric references (Shlesinger, 1994), idiosyncratic ellipsis (i.e., grammatical cohesive devices used as part of the interlocutors' shared knowledge (cf. Halliday & Hasan, 1976)), coded and telepathic language, and unintelligible items. These inherent peculiarities of EAR lead to calling into question the adequacy and consistency of the prevalent translation method, which thus motivates the present study.

2. Speech type and verbatim translation

The source language of EAR is talk-in-interaction, and thus bears all the stylistic hallmarks of natural speech, including the use of dialectal or colloquial words or expressions, abbreviated forms, ellipsis and prosodic features, as well as complementary non-verbal visual means such as accompanying physical gestures. In natural spoken language, parts of utterance and discourse meanings are communicated through style, non-verbal cues, and voice. According to Crystal (1997, p. 171), pitch and loudness are "the source of the main linguistic effects", which, along with effects "arising out of the distinctive use of speed and rhythm, are collectively known as the prosodic features of language".

In this paper, the source language is dialectal and generally characterised as informal/illiterate at the crossroad between colloquial Arabic (cf. Versteegh, 2001) and Modern Standard Arabic (MSA), the official written and formal spoken language in the Arab world. Arabic dialects are an analytical and simplified version of "the synthetic language system which was their starting point" (Holes, 1995, p. 157). They share many features, including lexical items, morphological patterns (e.g., verb patterns), negations, and word order, but differ in others, namely the phonological processes (particularly between the Western and Eastern dialects) and lexical and idiomatic usage. These differences evolved from sociological, cultural and historical circumstances, and they are not systematic (Holes, 1995). Phonological and lexical differences exist even between regional dialects of the one country, but are less problematic (Holes, 1995; Watson, 2002). Arab speakers of inter- and intra-dialectal varieties overcome their linguistic differences by switching to MSA or to each other's dialects (cf. Versteegh, 2001).

The range of conversational strategies available for a speaker is socially determined – "an individual's set of habitual strategies is unique within that range" (Tannen, 1982, p. 218) – and the speaker-listener cannot be idealised as belonging to a homogeneous language community (Foulkes & Docherty, 2006). Moreover, being familiar with a dialect minimises but does not rule out dialectal problems that might be encountered by EAR translators. Assuming that the translator is working within his/her range of dialectal expertise, the universal ethical rules of accuracy for translators and interpreters should apply when problems of comprehension are encountered due to idiolect (a person's individual speech pattern) or communal dialect (a community's speech pattern that is geographically, socially, culturally, and/or ethnically determined). This includes disqualifying oneself from the assignment or, if the problems are isolated, seeking assistance of colleagues or native speakers who are familiar with the language variety. With regard to speech production and reception, studies relating to between-speaker differences (e.g., Smirnova et al., 2007), between-listener differences (e.g., Grabe et al., 2005) and between-gender differences (e.g., Rosenhouse, 1998), within the same spoken variety, are work in progress but worthy of investigation in the context of the present topic.

Apart from its generic features and language-specific dialectal diversity, peculiar features of the speech type at hand have a bearing on the translation/transcription method and process. Listening device recordings are often of low quality compared with telephone interceptions. Because the conversation is private in the EAR material and all interlocutors – save any undercover agent, if involved – are unconscious of the recording of their conversation, the translator often lacks the necessary contextual knowledge for

interpretive decisions (Bucholtz, 2000). The purpose of the surveillance exercise is to probe into private life for prosecution, and with this end in view the police are on the lookout for criminal intention, conspiracy or confessions of commission of crime. The use of code-switching is a common feature of EAR, with languages other than English (LOTEs) being strategically mixed with English (which is often the speakers' first language) and sometimes Pig Latin (by juveniles, e.g., *abbingstay: stabbing*) if the interaction is taking place in an English-speaking country, to conceal messages or identity.

2.1 Addressing equivalence

To date, discourse-analysis research has rarely concentrated on prosodic features in interpreting (cf. Merlini & Favaron, 2005; Shlesinger, 1994), and, to my knowledge, never on the transcription of prosodic features in translated EAR. As previously discussed, these transcripts are conceived as prose 'written to be read' and not 'written to be read as if spoken'. Spoken and written languages are packaged differently (cf. the 'oral-written dichotomy' of Horowitz and Samuels (1987, p. 9)), and translating recorded material imposes the added responsibility of decision-making about the transfer of meaning using equivalent target language speech conventions (cf. Bucholtz, 2000). This transfer militates against the constraints of the 'verbatim' translation requirement, which rules that when 'interpreting' EAR – i.e. transcribing and translating it – "changing or paraphrasing of what is originally said is the same as altering testimony which is considered perjury or lying under oath" (Teichman, 2000, para. 3).

The production of an equivalent effect on the users of transcripts can be achieved through the faithful reconstruction of the interlocutors' participation in the formation of the message (cf. Nida, 1964). This requires making use of conversation analysis with reference to the prosodic and paralinguistic cues of speech to infer, translate and document meaning.

Selting and Couper-Kuhlen (1996) argue for the implication of prosody on meaning, and Gumperz points out "the shifts in intonation, volume, rhythm and tempo, that underlie prosodic assessments, and [explain] their grammatical functions" (1996, p. xi). He considers that prosody, among other indexical signs, including code- and style-switching, interacts "with symbolic (i.e., grammatical and lexical) signs, cultural and other relevant background knowledge (i.e., contextualization cues) to constitute social action" (as cited in Prevignano & di Luzio, 2003, p. 9) and mark, analytically, thematic coherence in various speech events (cf. Antonis et al., 2001; Halliday, 1994; Local, 1996).

Notwithstanding the hypothetical points of departure of the above interactional, linguistic and sociolinguistic approaches, the concern of this paper is the consensus on the integral function of prosody in framing meaning and its universal role in achieving "the pragmatic conditions of communicative tasks" (Gumperz & Gumperz, 1982, p. 12).

Time pressure, contemporaneousness and live interaction, which are the three pertinent factors that have driven research into spoken interpreting, are remotely relevant to the translation of EAR. Also, interpreting skills with regard to the prosodic features of spoken conversation have been given marginal attention compared with linguistic features, despite their impact on the output – be it 'spoken from spoken' material (cf. Shlesinger, 1994), 'spoken from written' material (cf. Agrifoglio, 2004) or 'written from spoken' material.

The genre and type of recorded voice conversation have not been tackled by translation studies, and the scarce discussions of translation 'techniques' of oral conversation (e.g. Edwards, 1995; Esposito, 2001) or transcription

techniques of monolingual EAR (e.g. Fraser, 2003), have not provided insight into the transfer of intention and actions, or about other issues, such as accuracy, clarity and naturalness.

In the absence of research on the topic in translation and interpreting studies, and given the relevance of prosody to the documentation of oral conversation, it is worth exploring the extension of application of pertinent notions in conversation analysis, in particular prosodic and paralinguistic cues, to the translation of EAR. The hypothesis is then that prosody in conversation is integral to the translation of EAR and failure to process and document prosodic features may result in misinterpretation (i.e. misinferencing) and mistranslation.

2.2 Prosody and meaning in transcripts

Gardner posits that, in conversation analysis, "the transcription process is the analysis" (1994, p. 103). He argues that, in order to obtain observable data, the process must not 'freeze' the interaction into a text, but rather 're-do' the event being analysed and 'create' "the meanings by the participants from moment to moment, and document the minute details of the interaction" (1994, p. 103). Such process revelation is axiomatic given that conversation involves roles and power relationships, exhibited in structurally organised and not chaotic exchanges, in spite of the presence of phenomena such as interruptions and simultaneous talk. In fact "nothing in conversation can be dismissed as disorderly, accidental [or] irrelevant" (Gardner, 1994, p. 102). Merlini and Favaron (2005) conducted an insightful investigation on power relationships and voice of interpreting in speech pathology, using prosody and conversation analysis. Their analysis shows the intrusive role that can be played by the interpreter who prosodically mismanages and manipulates the traditional doctor-patient interaction.

Examples of paralinguistic cues include *ums, mms, uhs* or *uh-huhs* in English, which can express hesitation, incomprehension and show that the interlocutors are listening or pretend to be listening to each other (cf. Sacks, 1995, pp. 746-747). Apart from its use as a way of 'sudden remembering' (Jefferson, as cited in Local, 1996, p. 178), *oh* can indicate a 'change-of-state token' (Local, 1996, p. 178) when its phonetic parameters are analysed with reference to lexis and syntax, e.g., displaying 'news-receipt' or 'partial repeats of prior turn' (Local, 1996, pp. 180-201). Other 'intervals and between utterances' (Gardner, 1994, pp. 185-186), i.e., pauses, and prosodic features, include coughing, laughter (indicating for example intimacy), and sound prolongation. These cues are relevant in the translation of EAR material, which also includes the important factor of inter-language transfer. A brief pause, for example, can indicate "a 'slot' which the referring speaker leaves open for the recipient to insert, in the case of success, a token of recognition" (Müller, 1996, p. 135). In this sense, translating EAR is largely influenced by the chosen transcription method.

The translator of other types of audio recordings (e.g. focus group) may have a licence to deliver pragmatic meanings in the prose, in parenthetical comments (e.g. "expressing anger") or in footnotes, without recourse to the transcription of prosodic features. This is a luxury the translator of EAR does not, and should not, have. It is noteworthy that this paper considers prosody as a universal property of language and is solely concerned with its interactional use. This includes the employment of the same prosodic feature for different functions, e.g., tune (O'Connor & Arnold, 1973), pause, latching, and different prosodic features for similar functions (e.g., the feedback tokens *yeah* and *okay*, sound prolongation).

2.2.2 Function of intonation. The prominence of prosodic features, such as intonation, to meaning is not new. As succinctly put by Shlesinger (1994, p. 225), "The functionality of intonational choices and their role in facilitating (or obstructing) communication is by now a universal point of departure in the literature."

Antonis et al. (2001) identify three notions around which the 'structural' functions of intonation are centred: *prominence* (provision of weight structuring of linguistic units, i.e., stress distinction on syllables), *grouping* (provision of coherence and segmentation of linguistic units into prosodic units, i.e., provision of tonal prominence or focus) and *discourse* (structuring prosodic units according to topics of discussion and turn regulations between speakers in a conversation). Halliday (1994) posits three intonation functions in spoken language: grammatical (relating to mood structure, e.g., question/statement, and identification of clause and sentence constituents), attitudinal (relating to speakers' attitudes and emotions, e.g., surprise, indifference, irony, gratefulness) and informational (indicating importance of a word and marking given and new information).

Couper-Khulen and Selting's (1996) interactional approach converges with the preceding functions of intonation outlined by Halliday (1994) and Antonis et al. (2001), but extends beyond these functions to embrace prosody in general. Crystal (1969) also argues that the meaning of intonation works in tandem with other variables, including, prosodic and paralinguistic systems, lexis, style, and kinetic cues. Prominence, grammatical, attitudinal, and discursive objectives inform then prosodical choices in speech. As they are conventional and acquired (cf. Cook, 2002), the tools with which these objectives are achieved are largely finite.

Bolinger also posits that intonation is located within the general scheme of iconic nonverbal communication and defines intonation as "primarily a symptom of how we feel about what we say, or how we feel when we say" (1989, p. 1). In a chapter dedicated to 'dialect and language', Bolinger (1989, pp. 26-64) argues for the universality of prosody, and in particular intonation across a variety of languages, including Syrian Arabic (cf. Syro-Lebanese dialectal classification in Versteegh, 2001). With reference to Arabic and English, Bolinger concludes that "the similarity between the two languages is remarkable, including certain rather small details" (1989, p. 54). Intonation is suprasegmental, and despite their segmental differences, Arabic and English are phonologically stress-timed. Moreover, like English, Arabic can take no more than one stressed syllable per word (cf. Holes, 1995, on stress in Modern Standard Arabic and Ghazali et al., 2002, on stress-timing variation in Arabic dialects). Apart from the consensus on the difference in intonation in tag and wh-questions between English and Arabic (cf. de Jong & Zawaydeh, 1999; Odisho, 2005), significant prosodic similarities between English and Arabic are reported by researchers other than Bolinger (e.g. Chahal, 1999; de Jong & Zawaydeh, 1999; Holes, 1995; Jun, 2005; Odisho, 2005).

The phonological, prosodic and intonational closeness between English and Syrian Arabic (the dialect of the speech excerpt in this paper) provides the ground for the use of approaches to the analysis of interactional and functional aspects of prosody and intonation described in the literature on the English language for inference, translation and transcription of EAR material. This is supported by the fact that the focus of this paper is on the emotions provided by prosodic features, rather than on the speaker and word recognition, accent identification, and topic and sentence segmentation. Furthermore, the processing of segmental information in speech is closely linked to stress.

Languages and dialects, as commonly held, differ markedly in the details of phonetic realisation of prosodic patterns (see Gibbon, 1998; Kulk et al.,

2003), but influential research on intonation functions and patterns (see Bolinger, 1978, 1989; Lieberman, 1967; Pierrehumbert & Hirschberg, 1990) is based on a universalist approach to intonation rather than a typological one (see discussion on both approaches in Ladd, 2001), or an approach which has gained universality through application (see Fujisaki, 1983; O'Connor & Arnold, 1973). O'Connor and Arnold (1973) made a pioneering study on the role of intonation in emotional expressions in colloquial English (middle-class southern British English). Their seminal work remains an emblematic resource for learning intonation in English teaching and in research on intonation (cf. Grabe et al., 2005; Gussenhoven, 2004). O'Connor and Arnold (1973) identify seven tunes, or pitch treatments, that are used to create intonation for syllables and words, and hence signal intentions and attitudes and provide a means of steering inferential processes.

They also systematically identify ten tone groups in spoken English, each of which expresses different sets of attitudes in different contexts and its tunes have one or more pitch feature in common. These tone groups are: low drop (expressing, for example, reservedness in statements or intensity in wh-questions); high drop (e.g. demanding agreement with question tags); take-off (e.g. appealing to the listener with commands); low bounce (e.g., encouraging effect with interjections); switchback (e.g. astonishment in echoed questions); long jump (e.g., protestation in statements); high bounce (e.g., tentativeness in straightforward wh-questions); jack-knife (e.g., antagonism with yes-no questions); high dive (e.g. pleading with commands); and terrace (e.g. calling out to someone with interjections).

O'Connor and Arnold's (1973) categorisation of 'tone groups' is appropriate for the analysis of attitudes and emotions conveyed by speakers in evidentiary audio recordings, as it is in tune with the interactional, linguistic and sociolinguistic approaches to the function of intonation as discussed above. The relevance of the authors' categorisation is also prompted by the speech pair used in the excerpt, which is the Syro-Lebanese variety of Arabic and Australian English, described by Gupta (1997) as 'monolingual ancestral English'.

3. Procedure and instrument

To test the above assumptions, an authentic Arabic speech excerpt from an evidentiary audio recording was faithfully transcribed, translated and analysed. The excerpt is from a speech of one hour and ten minutes, recorded by a crime squad in an Australian gaol using a listening device. The conversation is in Lebanese Arabic mixed with English, between two males charged with a criminal offence and awaiting a court hearing. The main criterion for choosing a conversation recorded through a listening device is to highlight the potential unintelligibility and incoherence that confront the transcriber, and discuss the relevance or otherwise of incorporating prosodic and paralinguistic information among the features of intelligible speech.

In discussing the meaning of representative examples, consideration will be given to interlocutors' style, monitoring, planning, and control in the conversation, with reference to their speech presentation on transcripts.

3.1 Transcription conventions
Crime agencies have opted for a minimalist approach to the documentation of prosodic features – restricting them, often cosmetically, to feedback, receipt tokens and hesitation (see Table 1 below). The transcription conventions need

to serve as an analytical instrument of the translated material and suit the purpose of producing a legible, but also defensible, transcript.

The transcripts of EAR are to be read by non-linguists and non-conversationalists, including judges, lawyers, police investigators, jury members, and the accused. According to Duranti (1997, p. 142) "the process of transcribing implies a process of socialization of our readers to particular transcribing needs and conventions". For the purposes of the present work, the transcribed translations need to be faithful to the original, read as fluently as possible, and at the same time be highly informative as to the prosodic and paralinguistic cues of the spoken utterances.

In conversation analysis, different scholars have used different transcription conventions to serve their research needs. Paul ten Have (1999) discusses the generic information available in monolingual transcripts, including archival information, such as time, date, and inferential information, such as overlapped speech and stresses. He and others (e.g., Edwards, 1995; Fraser, 1996) warn against the contentious issue of voice identification. The IPA transcription format is largely inaccessible, although it may have its use in computational and phonetic analysis. Substituting general rules of orthography to indicate phonetic features (e.g. 'dz' for 'does' and the vernacular 'yer' for 'your') is also, obviously, impractical and does not serve the purpose (Duranti, 1997). Accordingly, a hybrid transcription system, selected from ten Have (1999), Gumperz and Roberts (1991) and Gardner (1994), is adopted in the translation of the excerpts below. The selection aims at covering prosodic and paralinguistic cues pertinent to deciphering meaning, minimising visual disruption, and providing relative ease in word-processing.

Table 1. Transcription Conventions

Symbol	Significance
Simultaneous, overlapping and latched utterances	
[ ]	In *overlap*, the point of onset is marked with left hand square brackets, and the point at which overlap stops is marked with right-hand square brackets.
=	*Equal signs* link contiguous stretches of talk between which there is no gap and no overlap (*latched*).
= and [ ]	If more than one speaker latches onto a previous turn, this is shown through a combination of equal sign and square bracket.
Prosodic features of utterances	
(0.0)	*Extended pause- Numbers in parentheses* indicate elapsed time in silence by tenth of seconds, so (7.1) is a pause of 7 seconds and one tenth of a second. Silences can occur within utterances (*pauses*), or between utterances (*gaps*).
(.)	*Brief pause - A dot in parentheses* indicates a pause of less than 0.2 seconds within or between utterances.
word	*Underscoring* indicates some form of stress, via pitch and/or amplitude.
::	*Colons* indicate prolongation of the immediately prior sound. Multiple colons indicate a more prolonged sound.
-	*A dash* indicates a sudden cut-off (truncation) of the current sound.
. , ?	Punctuation marks are used to indicate characteristics of speech production, especially intonation; they do not refer to grammatical units.
.	A *full stop* indicates a falling terminal contour, i.e., a 'final' intonation.
,	A *comma* indicates a continuing intonation (contour) fall in tone, like when reading items from a list. It represents talk that is hearably incomplete.
?	A *question mark* indicates a strongly rising intonation (contour).

	Its characterising feature is that it rises a long way in pitch and ends up at the *high* end of the pitch range.
	The absence of an utterance-final marker indicates some sort of 'indeterminate' contour.
↑↓	*Arrows* indicate marked shifts into higher or lower pitch in the utterance part immediately following the arrow.
WORD	*Upper case* indicates especially loud sounds relative to the surrounding talk.
°	Utterances or utterance parts bracketed by *degree signs* are relatively quieter than the surrounding talk.
>words<	Right/left pointing *carets* bracketing an utterance or utterance-part indicate *speeding up*.
<words>	Left/right pointing *carets* bracketing an utterance or utterance-part indicate *slow talk*. It is optional as it can be apparent from colons indicating lengthening.
Feedback, receipt tokens and hesitation	
	Continuers and completers: okay, uh hm, mm hm, mm, oh, yeah, right, alright, you know *Hesitation marker*: uh, uhhh, for short or longer hesitation, respectively.
Translator's doubts and comments	
( )	*Empty parentheses* indicate inability to hear what was said. The length of the parenthesised space indicates the length of the untranscribed talk. In the speaker designated column, the empty parentheses indicate inability to identify the speaker.
(( ))	*Double parentheses* contain transcriber's/translator's description rather, or in addition to, transcriptions, i.e., vocalisations or non-linguistic vocal effects that indicate voice qualifications but cannot be satisfactorily transcribed in symbols.

3.2 Translation method

The constraints imposed on the translator by the outlined speech-type purpose and by the obligation of producing 'verbatim' transcripts imposed by crime agencies and the legal system, clearly demonstrate the appropriateness of Newmark's (1988) 'faithful translation' method for EAR material, as opposed to the word-for-word or literal methods (1988). Newmark defines faithful translation as a method that

> …attempts to reproduce the precise contextual meaning of the original within the constraints of the TL grammatical structures. It 'transfers' cultural words and preserves the degree of grammatical and lexical 'abnormality' (deviation from SL norms) in the translation. It attempts to be completely faithful to the intentions and the text-realisation of the SL writer (1988, p. 46).

Newmark refers here to the translation of texts, which are syntactically well organised and evenly punctuated. Natural speech, however, exhibits paratactic constructions (predominant use of coordination and juxtaposition) instead of the hypotactic constructions in writing (predominant use of subordination).

In *An Introduction to Functional Grammar*, Halliday identifies two different kinds of complexity associated with written and spoken languages as follows:

> Typically, written language becomes complex by being lexically dense: it packs a large number of lexical items into each clause; whereas spoken language becomes complex by being grammatically intricate: it builds up elaborate clause complexes out of parataxis and hypotaxis (1994, p. 350).

Halliday defines parataxis as "the relation between two like elements of equal status, one initiating and the other continuing, whilst hypotaxis is the relation between a dependent element and its dominant, the element on which it is dependent" (1994, p. 218). From the perspective of syntax, Newmark's definition of faithful translation is appropriate, but from the perspective of punctuation, it is inadequate, as will be discussed in the subsequent sections. In speech, pauses and stops are communicated phonologically and not graphically. Interestingly, Halliday (1987, p. 66) argues that spoken discourse is grammatically more intricate than written discourse, adding that "The more natural, un-self-monitored the discourse, the more intricate the grammatical patterns that can be woven." Earlier, Halliday and Hasan (1976, p. 5) state that "meaning is put into wording [i.e., words, grammar and syntax], and wording into sound or writing" (emphasis added). This aligns with the established fact that prosodic and paralinguistic features can form an integral part of the meanings of lexico-grammatical structures in the spoken mode.

Drawing on Halliday's work, Eggins (2004, p. 255) argues that when we talk we often chain clauses in sequence "and use markers to show the relationship between clauses (e.g., when, because), and we also use the spoken language systems of rhythm and intonation to signal to our listeners when we've reached the end of a clause sequence" or a clause complex. In systemic functional linguistics, a clause complex refers to the "grammatical and semantic unit formed when two or more clauses are linked together in certain systematic and meaningful ways" (Eggins, 2004, p. 255). Castello (2008) further argues that lexico-grammatical intricacy is reflected in how many clauses join together to form a clause complex.

It is this grammatical intricacy and complexity of spoken language which motivates and informs the translation method to be adopted for EAR material, with the aim to maintain the stylistic accuracy in line with the rules applicable to the treatment of register and stylistic variations in interpreting (Hale, 1997, 2002). One handy and accessible feature to account for style is the apostrophe (') to signal a missing sound expected in Standard English (cf. Labov as cited in Duranti, 1997), e.g. *'cause you ain't goin' to*, rather than the formal *because you are not going to*. Intercultural variability requiring the translator's intervention to achieve target language equivalence is linguistic and stylistic as discussed and exemplified.

4. Sample analysis

In the excerpt below, the interlocutors speak the North Lebanese dialect. The speech is significantly of poor sound quality due largely to noise, background conversation, echo, and, presumably, the unconsciousness of the interlocutors of the position of the listening device.

The conversation is mainly in Arabic, exhibiting now and then code-switching, borrowing of English words and phrases which are adapted to the grammatical rules of Arabic (Arabised). The social motivation for switching at the phrase level appears to be for convenience and fluency, e.g., يو ركن فيون يشوفو: '**you reckon** they can see', while need seems to be the driver behind the lexical switching, such as شرّج:'*to charge*'.

In the following analysis, use is optimally made of the texture, structure and prosodic features at hand, and reference is made to examples that highlight the importance of the transcription of prosody onto a stylistically faithful translation.

Table 2. Excerpt of translation of evidentiary audio recording

	MVI: male voice one; MV2: male voice two.	Texts in **bold** are spoken in English
1	(Noise and echo)	(Noise and echo)
2	MV1: ( )	( )
3	MV2: ( ) ↑did they **charge**( )?	↑**شرّجو**( ) ( )؟
4	MV1: ↑no (1.0) but he <showed( )>.	↑لأ (1.0) بس <أرجا( )>.
5	MV2: they're ↑goin' to **charge** ( ). ( ) (.) they're ↑goin' >to **charge** ( )<.	↑حي**شرّجو**( ). ( ) ↑رح> (.)ي**شرّجو** ( )>.
6	(whistle noise) (3.0)	(whistle noise) (3.0)
7	MV1: ( )	( )
8	( )	( )
9	MV2: ( )	( )
10	MV1: sixty ↑six	↑ستّ وستين
11	MV2: ↑are they goin' to **charge** ( )?	رح ي**شرّجو** ( )؟
12	MV1: huh? ( ) they showed me everything. (2.2) the photos,	هه؟ ( ) فرجوني كلشي. (2.2) الصور،
13	MV2: yeah.	إيه.
14	MV1: (1.0) our face.	(1.0) وشّنا.
15	( ): (° °)	( ) (° °)
16	MV1: ( ) our face, ( ).	( ) وشّنا، ( ).
17	(Unintelligible background conversation)	(Unintelligible background conversation)
18	(2.5)	(2.5)
19	MV1: but uh ( ) (1.5) in the **truck**. (1.0) ( )	بس أه ( ) (1.5) بال**تراك**. (1.0) ()
20	MV2: ( )=	=( )
21	MV1: =the three **jobs** I did in the car.	=التلات **جوبات** يللي عملتن بالسيارة.
22	( ) (1.0)	(1.0) ( )
23	MV1: the ↑photos. ( )(.) they've talked with( ). (.)	الـ↑صور. ( ) (.) حكيو مع ( ). (.)
24	MV2: ↓yeah.	↓إيه.
25	MV1: they've talked with him.	حاكيين معو.
26	MV2: ↑did he talk with( )?	↑حاكيين مع( )؟
27	MV1: ↑yeah.=	↑إيه.=
28	MV2: =↑what did he say?	=↑شو قال؟
29	MV1: **(.) I don't know.**	**(.) أي دونت نو.**
30	MV2: ( )	( )
31	(Noise, echo and unintelligible background conversation)	(Noise, echo and unintelligible background conversation)
32	(2.0)	(2.0)
33	MV1: the photos of the house, ( )	صور البيت، ( )
34	(2.8)	(2.8)
35	( )	( )
36	(unintelligible background conversation)	(unintelligible background conversation)
37	MV2: ( ) ↓**you reckon** >they can see ( )<?	( ) ↓ **يو ركن** >فيون يشوفو( )>؟
38	( )	( )
39	MV1: ↑our face?	↑وشّنا؟
40	(unintelligible background conversation)	(unintelligible background conversation)
41	(2.5)	(2.5)
42	MV1: **all of them.** ↑**all of them.**	**أول أوفم.** ↑ **أول أوفم.**
43	MV2: ↑**huh**?	↑هاه؟
44	(2.0)	(2.0)
45	( ): ( ) there are knuckledusters.	( ) في بونيات.
46	(unintelligible background conversation)	(unintelligible background conversation)
47	(1.5)	(1.5)
48	MV2: ( )	( )
49	MV2: there's the ↑**shirt**. Inside, it's ↑clear. (4.2)	في ↑القميص. جوا، ↑مبيّني. (4.2)
50	(unintelligible background conversation)	(unintelligible background conversation)

51	(1.5)	(1.5)
52	(unintelligible background conversation)	(unintelligible background conversation)
53 54	MV1: <u>but</u> (4.0) in the <u>photos</u> it's ( ), (.) other than the <u>shirt</u>, ( ) the <u>shirt</u>, ( ) >↑**you know what I mean?**<	بس ( ) بالصور هي ( )، (.) غير القميص، ( ) القميص، ( ) ↑**يو نو وات أي** <مين؟>
55	(unintelligible background conversation)	(unintelligible background conversation)
56	( )	( )
57	MV1:() there nothing except the ↓photos. he showed ( ) to me (.) **one hundred per cent.**	() ما في إلا الـلـصور. فرْجاني ( ) (.) وَن هندرد **برسنت.**
58	( ) (background noise and echo)	(background noise and echo) ( )
59	MV2: They're goin' to **charge** you, I heard.	رح يـ**شرّجـ**وك قال.
60	MV1: ( )?	( )؟
61	MV2: they're goin' to **charge** you (.) ( )	( ) (.) رح يـ**شرّجـ**وك
62	(2.5)	(2.5)
63	MV1: thirty one ( ).	واحد وثلاثين ( ).
64	MV2: ↑thirty one? (3.0) ↑serious.	واحد وثلاثين؟ (3.0) ↑لا.
65	MV1: ( ), ↑thirty one.	( )، ↑واحد وثلاثين.
66	MV2: ↑they're goin' to **charge** you with them ↑now?	رح ↑يـشرّجوك فيون ↑هلّق؟
67	MV1: **I don't know**, ( )	أي **دونت نو**، ( )
68	MV2: (.) they showed ( ), but they didn't tell ( )	(.) فرجو( )، بس ما قالو( )
69	( ) (background noise and echo)	(background noise and echo)( )
70	MV1: ( ) they're trying ( ) to <u>make</u> me ↑talk.	( ) عم يجربو ( ) حتى يخلّوني ↑إحكي.
71	( ) (background noise and echo)	(background noise and echo) ( )
72	MV1: ↑**you know what I mean**? ( )	↑**يو نو وات أي مين؟** ( )
73	( ) (background noise and echo)	(background noise and echo)( )

The meaning of the utterance in row 4 and the first utterance in row 5 is determined by the **low drop tone group** in both, which conveys a definite and complete opinion, marked by the absence of a nuclear head (i.e., the first stressed word or syllable in the tone group), hence expresses detachment (O'Connor & Arnold, 1973, pp. 47-48). By contrast, the second utterance in row 5 'they are ↑goin' >to <u>charge</u> ( )<' has two different features: (1) its tone group is **'high dive'** (O'Connor & Arnold, 1973, pp. 82-88), which also expresses unreservedness, emphatic definiteness and completeness through the **accented (stressed)** 'charge', and (2) the **accelerating tempo** exhibited through the '**clipped syllable**' (word) 'charge'. This corresponds with Crystal's findings that 'clipped syllables' frequently co-occur with a nuclear tone (here, 'charge'), "regardless of other pitch features co-occurring, which suggests that this feature is an important modifying factor in the interpretation of pitch glides" (1969, p. 154) . Further, based on an experiment involving the application of various feelings and intentions to statements by a number of subjects, Crystal (1969, p. 305) found that **fast tempo** is used to express conspiracy, among other uses, which is relevant to the topic at hand. This gives rise to the contrast in meaning, between the two adjacent and identical utterances, i.e., 'they will charge someone?' in the latter utterance as opposed to the former 'they may charge someone'. This suggests that the documentation of accent, intonation and tempo could be crucial.

Noteworthy also is the **unintelligibility** of the personal pronoun affixed to the verb endings in rows 3, 4 and 5, which in Arabic refers, grammatically, to the object. In rows 3 and 5, MV2 says, شرّجو ( ): 'they charged ( )'. Two possible objects can be inferred: 'you' or 'him'. The missing part of the last syllable 'you' شرجوك = *charra/juk*: 'they charged you', or 'him' شرّجوِه = *charra/juh*: 'they charged him', excludes the object 'me', given that no other

syllable is heard (compare شرجوني = *charra/ju/ni:* 'they charged me'). The verb '*charraja*' is a loan word from English. The last possibility is also inconsistent with the context and legal process, knowing that a person is told when they are charged. Guesswork can be detrimental here, so the use of ( ), i.e., unintelligible, is important, rather than the arbitrary translator's/ transcriber's comment of 'sounds like'.

In row 4, the co-text (the adjacency pair, question/answer) combined with the **decelerated tempo** of ( ) أرجا: he showed ( ), suggest two possible 'objects': 'me' or 'him'.

In row 15, the **degree marks** in the unintelligible brackets imply whisper, the meaning of which can be important in other contexts as it may evoke conspiratorial intention.

In row 24, the **falling tone** in '*yeah*' communicates an invitation to the speaker to 'continue' as in *mm hm*.

The **latching** in rows 27 and 28 expresses MV2's concern in this context, but could equally express power in a larger context.

The pronoun in row 37 is not clear; it sounds like 'us' or 'me'. However, the following clarifying question, 'our face?' indicates the likelihood of 'us'. The **accelerated tempo** in row 37 conveys impatience and irritability (cf. Crystal, 1969, p. 305) about the fact that 'they can be identified from the photos'.

In row 39, the noun-pronoun 'our face' disagreement in number is common in spoken Arabic and exhibits, through the faithful translation method adopted, informal colloquial usage but not the speaker's educational level in Arabic.

The 2.5 second **pause** in row 41 helps in delivering the message, given that the gap between the clarifying rhetorical question in row 39 (also expressed through intonation), 'do you mean our faces?/you mean our faces?' and the second heard utterance by the same speaker (row 42) is relatively short and interrupted by something said. Hence, the utterances are clearly an emphatic answer to the question asked originally by MV2 in row 37.

In row 59, رح يشرجوك قال literally means 'they're going to charge you, he said'; however, in the absence of a referent, the end position of 'he said' in this instance, imparts 'I heard'. This is also a **topic shift/opener** strategy in Levantine (also known as Greater Syria, including, linguistically, Lebanon) Arabic conversation, expressing knowledge that the listener already has.

The **low drop tone group** in row 61 conveys a detached statement and not a question. The above discussion demonstrates that prosodic features can be very crucial in certain instances, to infer meaning despite the absence of linguistic information.

In sum, the faithful translation method applied to the excerpt above would have generated serious comprehension problems to the reader/analyst without recourse to prosodic features. To put this 'uncompromising and dogmatic' translation method on a par with the 'flexible' semantic translation (Newmark, 1988, p. 46), use was made of available prosodic features, among other strategies. Therefore, Newmark's (1988, p. 46) definition of faithful translation of EAR calls for the following extension:

Translating evidentiary audio recordings must be conducted through a faithful translation method that attempts to reproduce the contextual meaning of the original *speech* within the constraints of the TL grammatical structures. It transfers cultural words and preserves the degree of grammatical and lexical abnormality (deviation from SL norms). It attempts to be completely faithful to the intentions and the *speech-realisation of the SL speaker via linguistic, paralinguistic and prosodic means as necessary.*

Table 3 below illustrates some significant strategies used to optimally manage each component of the faithful translation rationale. The right hand column shows possible erroneous interpretation in the absence of prosodic and/or paralinguistic cues.

Table 3: Intentions, attitudes or feelings marked by prosodic and paralinguistic cues

Example	Speech strategy of prosodic or paralinguistic feature/s	Intention marked by prosodic and paralinguistic feature/s	Possible erroneous interpretation without added prosodic and paralinguistic feature/s
they're goin' >to **charge** ( )<. (Row 5)	*High dive* intonation, *Fast tempo and accentuation*	Unreserved and emphatic definite statement.	They're goin' to **charge** ( ). - with a *jack-knife* intonation (*rise-fall tune ending*): Disclaiming responsibility - with a *switchback* intonation (*fall-rise* tune ending): Concerned.
They're goin' to charge you, I heard. (Row 59)	*A falling terminal tone group on tag statement (full stop) with absence of referent.*	A topic shift/opener device inviting the listener to chat about shared information.	"they're goin' to charge you, he said." - with a *switchback* intonation (*fall-rise on tag statement*): Questioning with a tone of surprise.
Managing intentions marked with non-intonational devices			
Example	Speech strategy/ turn taking/doubts /comments	Contextual meaning	Possible literal meaning
no (1.0) but he <showed ( ) (unintelligible syllable)>. (Row 4)	*Unintelligible syllable (as opposed to word)*	Marking unintelligible inflected pronoun. In the context, either 'me' or 'him'.	No one-to-one correspondence.
MV1: yeah.= MV2: = what did he say? (Rows 27-28)	*Latching*	Conveying concern.	Exhibiting power.

Table 3 provides a snapshot of intentions 'concluded' with reference to prosodic features. These cues, in tandem with the lexico-grammatical organisation, influence, and are influenced by the context and moods at the utterance, sequence and speech levels. Although a limited number of utterances are analysed, the discussion provides plausible hints about the overall mood of the excerpt: the interlocutors' concern about their situation and admission of involvement in a joint criminal enterprise.

The analysis has identified stretches of texts to which the prosodic analysis has no bearing, which serves to minimise distraction in the transcription and reading processes. Although the transcription conventions are applied indiscriminately in documenting crucial intended messages, they can be selectively used where they have significant relevance to meaning. According to ten Have (1999, p. 78), transcriptions "are always and necessarily selective." Analysts have traditionally devised interactional features transcribed to document details to texts written in standard orthography, according to their interest and need. The same principle applies to the transcription of EAR.

5. Putting theory into practice

The above discussion suggests that EAR translators must have the necessary knowledge and knowhow to account for the linguistic and stylistic features of their working dialects, and be aware of the interactional role of prosody (and

its transcription) in these dialects. Only trained translators who are native speakers of the source language dialect can meet these criteria.

The discussion also suggests that analysts in crime agencies in charge of analysing EAR translations for prosecution purposes need to have specialised training in conversation analysis. It is equally desirable for legal professionals to become familiar with the rationale and significance of prosody in transcripts and of the outlined translation method.

Translation and interpreting training programs, especially those encompassing legal interpreting and specialised translation, ought to include training in decoding and encoding prosody and paralinguistic cues in speech. The theory and practice of conversation analysis should be integral to interpreting courses at all levels, and to translation courses at an advanced level. In-house workshops are the obvious training forum for translators practising in the field and clients interested in EAR translations, particularly police investigators and analysts.

6. Conclusion

Many EAR translations become crucial pieces of evidence in courts. The success of pleadings and the futures of those under investigation can hinge on the precision of the translation and transcription. These tasks are usually performed from recorded audio conversations that lack visual aids such as non-verbal expression (looks or gestures) and the benefit of feedback that face-to-face interpreters enjoy. The occasional poor sound quality, the privacy of conversation, the regional and idiosyncratic dialectal problems, and the restrictions imposed on the provision of 'processed' translation, all lead to the production of texts which are tantamount to unintelligible words and sentences. The adoption of a minimalist approach to account for non-linguistic cues in the translation of conversations does not serve the purpose. Face-to-face interpreters (should) rely on, and (should) relay prosodic features, in particular intonation, to determine and transmit intentions, attitudes and feelings. The same process should apply to the translation of EAR. The literature on conversation analysis, prosody and intonation provides insights into the structure and realisation of meaning of the spoken language, the material in question. These studies have developed transcription conventions that can be employed to complete the puzzle of the simplistic and risky approach of 'faithfully' translating and transcribing speeches through standard orthography only. Here an evidentiary audio recording excerpt was comprehensively transcribed and faithfully translated and using conventional transcription symbols. The analysis proves that a number of these symbols are essential for the reader to infer the full meaning and mood of utterances, turn-takings, and ultimately the speech itself, and hence their incorporation in the transcript is necessary.

In the context of EAR translation, the interpretation, or reception, of prosodic features reproduced in English – the target language in this paper - rests with the crime agency's analysts, who must also match the translator's expertise in terms of use and usage of prosody and conversation analysis.

Pedagogically, the appreciation of conversation strategies and prosodic features, in addition to the framing and delivery of meaning, can help enhance the cognitive competence required by trainee interpreters for the processing and storage of information, and also assist in note-making and delivery.

Although the above claims hinge on the analysis of a short sample with reference to one speech pair, it is assumed that the translational, transcriptional and analytical methodology and instrument used could yield

similar results if applied to other speech pairs. Further investigation into the translation and transcription of conversation between interlocutors speaking different dialects or interlocutors whose dialects are different to that of the translator should provide overdue insight to those who are directly or indirectly concerned. Also, further empirical examination of the impact of the approach adopted here on the users of transcripts, including the translator as an 'expert earwitness', would demonstrate the relevance of the procedure to the justice system.

References

Agrifoglio, M. (2004). Sight translation and interpreting: a comparative analysis of constraints and failures. *Interpreting,* 6(1), 43-67.

Antonis, B., Granstström, B., & Möbius, B. (2001). Developments and paradigms in intonation research. *Speech Communication, 33*, 263-296.

Berk-Seligson, S. (2002). *The bilingual courtroom* (2nd ed.). London, Chicago: The University of Chicago.

Bolinger, D. (1978). Intonation across languages. In: J.P. Greenberg, C.A. Ferguson & E.A. Moravcsik (Eds.), *Universals of human language, Vol. II Phonology* (pp. 471-524). Stanford: Stanford University Press.

Bolinger, D. (1989). *Intonation and its uses: Melody in grammar and discourse.* London, Melbourne, Auckland: Edward Arnold.

Bucholtz, M. (2000). The politics of transcription. *Journal of Pragmatics, 32*, 1439-1465.

Castello, E. (2008). *Text complexity and reading comprehension tests*. Bern: Peter Lang.

Chahal, D. (1999). A preliminary analysis of Lebanese Arabic intonation. In *Proceedings of the 1999 Conference of the Australian Linguistic Society*. Retrieved from http://www.als.asn.au/proceedings/als1999/chahal.pdf

Cook, N. (2002). *Tone of voice and mind*. Amsterdam and Philadelphia: John Benjamins.

Couper-Khulen, E., & Selting, M. (1996). Introduction. In E. Couper-Khulen & M. Selting (Eds.), *Prosody in conversation: Interactional studies* (pp. 1-10). New York: Cambridge University Press.

Crystal, D. (1969). *Prosodic systems and intonation in English*. London and New York: Cambridge University Press.

Crystal, D. (1997). *A dictionary of linguistics and phonetics (*4th ed.). Cambridge, MA: Blackwell.

De Jong, K., & Zawaydeh, B.A. (1999). Stress, duration, and intonation in Arabic word-level prosody. *Journal of Phonetics, 27*, 3-22.

Duranti, A. (1997). *Linguistic Anthropology*. Cambridge: Cambridge University Press.

Edwards, A.B. (1995). *The practice of court interpreting.* Amsterdam and Philadelphia: John Benjamins.

Eggins, S. (2004). *An introduction to systemic functional linguistics* (2nd ed.). New York and London: Continuum.

Esposito, N. (2001). From meaning to meaning: the influence of translation techniques on non-English focus group research. *Qualitative Health Research, 11*(4), 568-579.

Foulkes, P. & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics, 34*, 409-438.

Fraser, H. (1996). Identifying evidentiary audio voices – what phonetic science can and can't do. *Policing Issues and Practice Journal*, *4*(4), 39-43.

Fraser, H. (2003). Issues in transcription: factors affecting the reliability of transcripts as evidence in legal cases. *Forensic Linguistics*,*10*(2), 203-226.

Fujisaki. H. (1983). Dynamic characteristics of voice fundamental frequency in speech and singing. In P. MacNeilage (Ed.), *The production of speech* (pp. 39-55). New York: Springer.

Gardner, R. (1994). Conversation analysis: Some thoughts on its applicability to applied linguistics. In R. Gardner (Ed.), *Spoken Interaction Studies in Australia. Australian Review of Applied Linguistics, Series S, 11*, (97-118).

Ghazali, S., Hamdi, R., & Barkat, M. (2002). Speech rhythm variation in Arabic dialects. In *Proceedings of Speech Prosody, Aix-en-Provence, 2002* (pp. 331-334). Retrieved from http://sprosig.isle.illinois.edu/sp2002/pdf/ghazali-hamdi-barkat.pdf

Gibbon, D. (1998). Intonation in German. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 78-95). Cambridge: Cambridge University Press.

Grabe, E., Kochanski, G., & Coleman, J. (2005). The intonation of native accent varieties in the British Isles - potential for miscommunication? In K. Dziubalska-Kolaczyk & J. Przedlacka (Eds.), *English pronunciation models: a changing scene* (pp. 311-337). Frankfurt am Main: Peter Lang.

Gumperz, J., & Gumperz. J.C. (1982). Introduction: language and the communication of social identity. In J.J. Gumperz (Ed.), *Language and social identity* (pp. 1-21). Cambridge: Cambridge University Press.

Gumperz, J.J. (1996). Foreword. In E. Couper-Khulen & M. Selting (Eds.), *Prosody in conversation: Interactional studies* (pp. x-xii). New York: Cambridge University Press.

Gumperz, J.J., & Roberts, C. (1991). Understanding the intercultural encounters. In J. Blommaert & J. Verschueren (Eds.), *The pragmatics of intercultural and international communications* (pp. 51-91). Amsterdam and Philadelphia: John Benjamins.

Gupta, A.F. (1997). Colonisation, migration and functions of English. In E.W. Schneider (Ed.), *Englishes around the World 1: General studies, British Isles, North America studies in honour of Manfred Görlach* (pp. 47-58). Amsterdam and Philadelphia: John Benjamins.

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Hale, S. (1996). Pragmatic considerations in court interpreting. *Australian Review of Applied Linguistics, 19*(1), 61-72.

Hale, S. (1997). The treatment of register in court interpreting. *The Translator: Studies in Intercultural Communication, 3*(1), 39-54.

Hale, S. (2002). How faithfully do court interpreters render the style of non-English speaking witnesses' testimonies? A data-based study of Spanish-English bilingual proceedings. *Discourse Studies, 4*(1), 25-47.

Halliday, M.A.K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Halliday, M.A.K. (1987). Spoken and written modes of meaning. In R. Horowitz & J. Samuels (Eds.), *Comprehending oral and written language* (pp. 55-82). San Diego and London: Academic Press.

Halliday, M.A.K. (1994). *An introduction to functional grammar* (2nd ed.). London: Edward Arnold.

Holes, C. (1995). *Modern Arabic: Structures, functions and varieties*. London and New York: Longman.

Horowitz, R., & Samuels, J. (1987). Comprehending oral and written language: critical contrasts for literacy and schooling. In R. Horowitz & J. Samuels (Eds.), *Comprehending oral and written language* (pp. 1-52). San Diego and London: Academic Press.

House, J. (2001). Translation quality assessment: Linguistic description versus social evaluation. *Meta XLVI*(2), 243-257.

Jun, S-A. (2005). Prosodic typology. In S-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 430-458). Oxford: Oxford University Press.

Kulk, F., C. Odé, & Woidich, M. (2003). The intonation of colloquial damascene Arabic: a pilot study. *Proceedings* 25, *the Institute of Phonetic Sciences of the University of Amsterdam*, 15-20.

Ladd, D.R. (2001). Intonation. In M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (Eds.), *Language typology and language universals: an international handbook* (pp. 1380-1390). Berlin and New York: de Gruyter.

Lieberman, P. (1967). *Intonation, perception, and language*. Cambridge, MA: MIT Press.

Local, J. (1996). Conversational phonetics: some aspects of news receipts in everyday talk. In E. Couper-Khulen & M. Selting (Eds.), *Prosody in Conversation: Interactional studies* (pp. 177-230). New York: Cambridge University Press.

Merlini, R., & Favaron, R. (2005). The voice of interpreting in speech pathology. *Interpreting 7*(2), 263-302.

Müller, E.F. (1996). Affiliating and disaffiliating with continuers: prosodic aspects of recipiency. In E. Couper-Khulen & M. Selting (Eds.), *Prosody in conversation: Interactional studies* (pp. 131-176). New York: Cambridge University Press.

Newmark, P. (1988). *A Textbook of Translation*. UK: Prentice Hall International.

Nida, E. (1964). *Toward a Science of Translation*. Leiden: E. J. Brill.

Odisho, E. (2005). *Techniques of teaching comparative pronunciation in Arabic and English*. New Jersey: Gorgias Press.

O'Connor, J.D., & Arnold, G.F. (1973). *Intonation of Colloquial English: A practical handbook*. London: Longman.

Pierrehumbert, J., & Hirschberg. J. (1990). The meaning of intonation contours in the interpretation of discourse. In P.R. Cohen, J. Morgan, & M.E. Pollack (Eds.), *Intentions in communication* (pp. 271-311) Cambridge, MA: MIT Press.

Prevignano, C.L., & di Luzio, A. (2003). A discussion with John J. Gumperz. In S. L. Eerdmans, C. L. Prevignano, & P.J. Thibault (Eds.), *Language and interaction: Discussion with John J. Gumperz* (pp. 7-29). Amsterdam and Philadelphia: John Benjamins.

Rosenhouse, J. (1998). Women's speech and language variation in Arabic dialects. *Al-'Arabiyya 31*, 123-152.

Sacks, H. (1995). Paraphrasing; alternative temporal references; approximate and precise numbers; laughter; 'uh huh'. In G. Jefferson (Ed.), *Lectures on Conversation: Volumes I and II* (pp. 739-747). Oxford and Cambridge, MA: Balckwell.

Shlesinger, M. (1994). Intonation in the production and perception of simultaneous interpretation. In S. Lambert & B. Moser-Mercer (Eds.), *Bridging the gap: Empirical research in simultaneous interpretation* (pp. 225-236). Amsterdam and Philadelphia: John Benjamins.

Smirnova, N., Starshinov, A., Oparin, I., & Goloshchapova, T. (2007). Speaker identification using selective comparison of pitch contour parameters. *16th International Congress of Phonetic Sciences, 6-10 August, Saarbrucken, Germany*. Retrieved from http://icphs2007.de/conference/Papers/1138/1138.pdf

Tannen, D. (1982). Ethnic style in male-female conversation. In Gumperz, J. (Ed.), *Language and social interation* (pp. 217-231). Cambridge: Cambridge University Press.

Ten Have, P. (1999). *Doing conversation analysis, a practical guide*. London: Sage.

Teichman, D.E. (2000). Interpreting evidentiary tape recordings: the toughest job you'll ever love, or maybe not. Retrieved from http://www.linguisticworld.com/documents/PDF%206%20ATA%20Chronicle%20article%20scan0001.pdf

Versteegh, K. (2001). *The Arabic language*. Edinburgh: Edinburgh University Press.

Watson, J.C.E. (2002). *The phonology and morphology of Arabic.* New York: Oxford University Press.