Edinburgh Research Explorer

Prosodic transcription of Glasgow English: an evaluation study of GlaToBI

Citation for published version:
Mayo, C, Aylett, M & Ladd, DR 1997, Prosodic transcription of Glasgow English: an evaluation study of GlaToBI. in *Intonation: Theory, Models and Applications.*

Link:
[Link to publication record in Edinburgh Research Explorer](#)

Document Version:
Peer reviewed version

Published In:
Intonation: Theory, Models and Applications

PROSODIC TRANSCRIPTION OF GLASGOW ENGLISH: AN EVALUATION STUDY OF GLATOBI

Catherine Mayo *, **Matthew Aylett, D. Robert Ladd**
University of Edinburgh,
Department of Linguistics
catherin@ling.ed.ac.uk, matthewa@cogsci.ed.ac.uk, bob@ling.ed.ac.uk

ABSTRACT

GlaToBI, a version of the ToBI prosodic transcription system which can be used to transcribe the intonation patterns of western Scottish (Glasgow) English, is currently under development. An assessment of GlaToBI, similar to the evaluation studies that were undertaken for the original ToBI system [7], and for GToBI, a version developed for German [4], has been carried out to test the new system's reliability, learnability and comprehensiveness. The results of this study show that this adaptation of the ToBI system can be applied with the expected level of reliability to the transcription of Glasgow English.

1. INTRODUCTION

Very little corpus based work has been done on the prosodic features of English dialects other than Standard American and southern British (Received Pronunciation). However, with the creation of databases such as the University of Edinburgh's HCRC Map Task corpus [1], the predominant dialect of which is western Scottish (Glasgow) English, the opportunity has arisen for investigation into the prosody of these minority dialects. Unfortunately no agreed upon system exists for the prosodic analysis of any non–standard English dialects: the original ToBI labelling system was developed specifically for the transcription of the "three most widely used dialects of English—namely general American, standard Australian and southern British" [2, p. 2], and is not necessarily adequate for other dialectal variants without adaptation. This paper is a brief report on GlaToBI, an adaptation of the ToBI system for Glasgow English.

1.1. Glasgow English intonation

Glasgow English, along with other urban north British dialects, differs from Standard American and RP English in its more frequent use of nuclear rises [3]. In addition, the Glasgow rise is often followed by what Cruttenden calls a "plateau–slump": after the pitch has risen at the accented syllable, it remains high (or slightly declining) until very near the edge of the phrase, at which point it falls again. In an analysis of the alignment of the high and low targets of Glasgow rises, Ladd [5] notes that the

typical rise in Glasgow English is a glide up on the accented syllable, normally beginning before and peaking after the accented syllable itself, a claim based on preliminary analyses of the HCRC Map Task corpus. These two elements—the rise and the following plateau–slump—are the points of focus for adaptation of ToBI for Glasgow English.
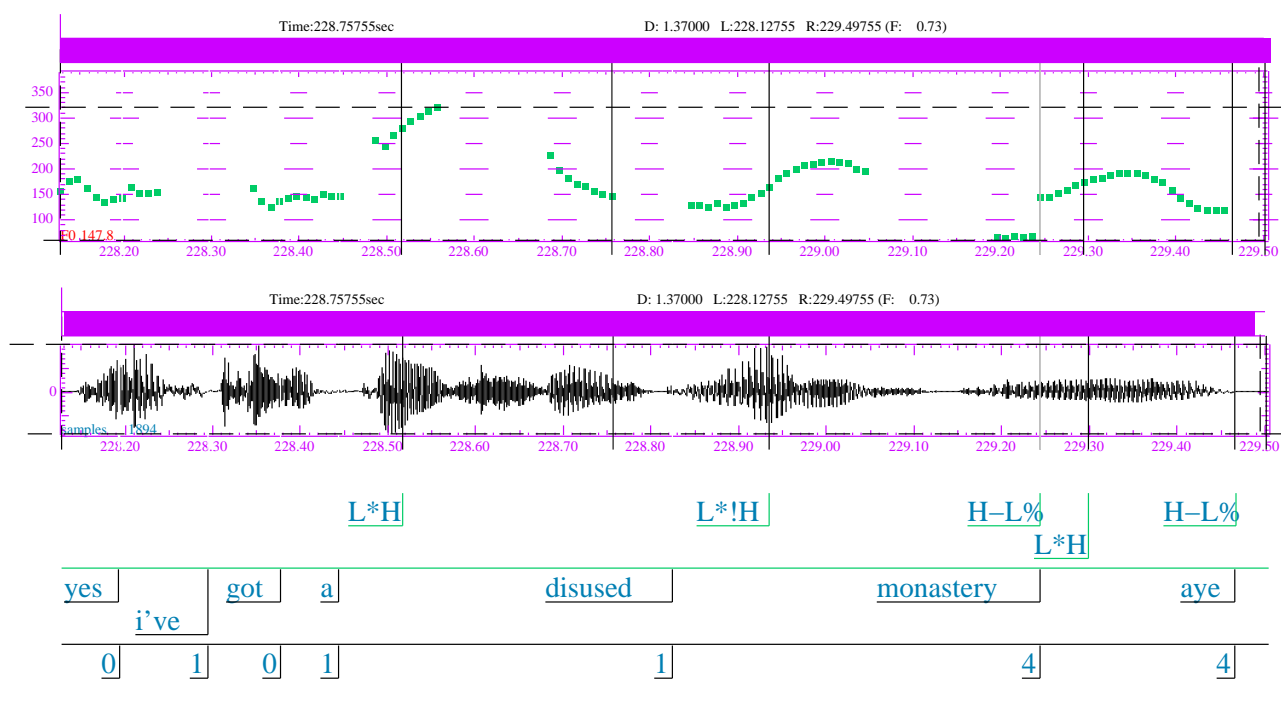
1.2. GlaToBI

As in the original ToBI system, intonation is transcribed in GlaToBI as a sequence of discrete pitch prominances (pitch accents) and boundary pitch movements (edge tones). In addition, edge tones, as in the original ToBI, are divided into two further categories: phrase accents, which mark the end of intermediate phrases, and boundary tones, which the mark the end of full intonational phrases. Following the ToBI system, which was based partly on work by Pierrehumbert [6], GlaToBI uses two main tones, high (**H**) and low (**L**), which can be combined in various ways to mark all of these intonational events.

The original ToBI system has two bitonal rising pitch accents: **L+H***, a step up from an unaccented syllable to the following accented syllable, and **L*+H**, a step up from the accented syllable to the following unaccented one. However, neither of these accents adequately describes the Glasgow English rise. In GlaToBI, therefore, both of the original ToBI rises have been removed and replaced by just one compound pitch accent **L*H**. The placement of the "*" diacritic between the L and H tones indicates that the stressed syllable does not align with either one or the other, but with the movement *from* one *to* the other.

The **H–** phrase accent in the original ToBI system triggers up–step, i.e. it raises the pitch of all following boundary tones. As a result, in ToBI the phrase–final compound tone **H–L%** is realised as a level tone relatively high in pitch. In order to accommodate the Glasgow English plateau–slump, up–step has been eliminated from the tonal inventory: H–L% in GlaToBI therefore represents a phrase final high–low sequence.

In addition to the above, GlaToBI has retained the downstep "**!**" diacritic to mark the overall decline of the speaker's fundamental frequency across an utterance as in the original ToBI system. The break index tier, which allows the transcriber to represent the strength of boundaries or breaks between words in an utterance on a scale

---

1. Glasgow English intonation contour with GlaToBI labels

of 0–4, has also been retained unchanged from the original system. The tonal inventory of GlaToBI is as follows:

Pitch accents:		Edge tones:	
H*		H–	
L*		L–	
L*H		H–L%	
!H*		H–H%	
L*!H		L–H%	
		L–L%	

2. EVALUATION STUDY

This section describes an evaluation study of GlaToBI system based on the goals set out for the original ToBI system: reliability, learnabilty and comprehensiveness [8]

2.1. Subjects

The subjects for this study were three of the four designers of the GlaToBI system ("experts"), and four transcribers who had not previously been exposed to the system ("non–experts"). The non–experts were all at least somewhat familiar with speech–related research, but only one had had some (rather limited) experience with the original ToBI system. All subjects were native speakers of English, with a number of different dialects represented, including one native speaker of Glasgow English.

2.2. Speech materials

The materials used for the evaluation were taken from the HCRC Map Task corpus. This corpus consists of task–oriented dialogues in which the participants have to reproduce a specified route on a map. The dialogues are therefore predominantly spontaneous speech, with some words occasionally in citation form (e.g. "Do you have

'carpenter's cottage' [on your map]?"). In the evaluation study, the subjects transcribed 30 utterances, comprised of 273 orthographic words. The length of each utterance ranged from 2 words to 19 words, averaging approximately 9 words per utterance.

2.3. Training

The four non–expert transcribers were given an intensive one day training session to familiarise them with ToBI–style transcription, and with GlaToBI, before participating in the evaluation proper. The example utterances used in the session were taken both from the training material provided with the original ToBI system (i.e. non–Glasgow English speech) and from the Map Task corpus itself. At the end of the training a pre–test was given to make sure that the system had been fully understood (a fifth non–expert was dropped from the study at this stage).

2.4. Evaluation proper

All subjects (non–expert and expert) were tested individually within the week following the training session. During the test, the subjects were not permitted to consult with each other or with the experimenter, but the non–experts were allowed to consult their training notes. No time limit was put on the test; most subjects finished transcribing 30 utterances in one day (approximately 8 hours in total, including breaks). The utterances and transcription tiers were presented to the subjects using Entropics *Waves+* speech analysis software in a UNIX environment (i.e. the standard ToBI platform).

2.5. Analysis

As in the evaluation studies carried out for the original ToBI [7] and GToBI [4], a pairwise analysis was used to

assess the level of agreement reached by the subjects in this study. This method, instead of comparing the labels of individual transcribers against the group, compares the labels of each transcriber against the labels of every other transcriber, for that particular aspect of the utterance. For instance, if four out of five transcribers label a word H*, and the fifth labels the same word L*H, the level of agreement is not considered to be 4 out of 5, but 6 out of 10— i.e. 6 pairs out of a possible 10 pairs of transcribers who agree as to the exact label. Pitrelli et al. [7] state that this is a stringent method of analysis, because it does not return the deceptively high 80% agreement which would be found by an analysis which simply stated that 4 out of 5 transcribers agreed.

3. RESULTS

273 words were transcribed by all seven subjects, totalling 1911 individual transcriptions, and 5733 transcriber–pair–words (273 words times 21 transcriber pairs). In the results reported here, as in the ToBI and GToBI studies, downstep was ignored for the purposes of analysis (e.g. !H* was considered to be the same as H*). Similarly, labels with the 'p' disfluency diacritic and the '–' uncertainty diacritic were also merged with their unmarked labels.

3.1. Pitch accents

In the analysis of agreement for pitch accents, the following elements were examined for each word in all utterances: (i) the rate of agreement as to the presence or lack of an accent on the word, and (ii) the rate of agreement as to the accent marked. The possible labels available for marking pitch accents were: H* (and !H*), L*, L*H (and L*!H), and no accent. The level of overall agreement as to whether a pitch accent was present on a word was 78% , while the overall rate of agreement as to which pitch accent was present was 68%.

Note that in this study, the analysis of agreement as to which accent/tone is marked on a word includes "no accent/tone" as a valid category. Given a situation in which two transcribers from a group of three or more agree that a word should be given a certain label, with the rest of the group leaving the word unlabelled, an analysis which examines agreement as to choice of accents/tones only when there is agreement that an accent/tone is present, will find that the rate of agreement for the above situation is 100% (i.e. all transcribers who marked an accent/tone agree to its identity). However, in a group of three or more transcribers, the pair who marked the accents are only one pair from a possible 6 or more: an agreement of 100% is therefore rather a false positive.

3.2. Edge tones

Edge tones were analysed in terms of (i) the rate of agreement as to the presence or lack of tone on each word, (ii) the rate of agreement as to the tone marked, and (iii) the rate of agreement as to the strength of the boundary marked (full boundary, marked by a boundary tone, intermediate boundary, marked by a phrase accent, or no boundary). For (ii) the labels available to the transcribers were: H– (and !H–), L–, H–H% (and !H–H%), H–L% (and !H–L%), L–L%, L–H%, and no boundary. Overall, the level of agreement as to the presence or lack

of edge tone was 92%, the highest level of agreement achieved in the study. The rate of agreement as to the edge tone marked was 82%, while transcribers agreed upon the strength of the boundary 88% of the time.

3.3. Break Indices

As it is technically not an option for no break index to be placed on a word, break indices were only analysed in terms of the level of break index marked. A small number of break indices were, however, acidentally omitted by transcribers (the software provided for the original ToBI evaluation which flagged illegal transcriptions was not available for the current study). For purpose of analysis, therefore, the categories of break index included a null category: 0, 1, 2, 3, 4, and no break index. The overall rate of agreement as to the choice of break index used was 74%.

3.4. Non–experts compared to experts

The use of non–expert as well as expert GlaToBI transcribers was critical in order to assess the ease and accuracy with which the new system could be learnt. As shown in Table 1, the level of agreement among the non–experts (N–E) was generally equal to or only slightly lower than among the experts (E), indicating that with the appropriate training, GlaToBI can be used by non–expert transcribers to a high level of consistency.

	N–E (%)	E (%)
presence of pitch accent	75	81
choice of pitch accent	62	69
presence of edge tone	92	93
choice of edge tone	82	82
strength of edge tone	88	88
choice of break index	77	72

Table 1. Rate of agreement for non–experts compared to experts

3.5. Results compared with ToBI and GToBI

The rates of agreement for GlaToBI compared to ToBI and GToBI are shown below in Table 2.

	GlaToBI (%)		ToBI (%)	GToBI (%)
presence of pitch accent	78		80.6	87
choice of pitch accent	65		72.4	74
presence of edge tone	92	(phrase accent) (boundary tone)	89.8 93.4	n/a
choice of edge tone	82	(phrase accent) (boundary tone)	85.3 90.0	86
strength of edge tone	88		n/a	86
choice of break index	74		70.4	n/a

Table 2. GlaToBI compared to ToBI and GToBI

3.6. Evaluation of L*H

The most fundamental change made to the original ToBI system was the replacement of the two original ToBI rises by L*H. Of the 273 words transcribed in this study, 105 were labelled L*H by at least one transcriber. When L*H is used unanimously (or confused only with "no accent") it tends to be in the nuclear position in the utterance, and followed by an H–L% tone—i.e. the prototypical Glasgow English rise–plateau–slump. However, 90 of the 105 words labelled L*H in this study were also marked with a different pitch accent by at least one other transcriber. 40% of these 90 were instances where L*H was also labelled with L* and with H*, while 43% were disagreements between L*H and H* only. Ladd [5] notes that for Glasgow English speech "the alignment and scaling of the L* and H tones appear to be quite variable. It is not clear whether all this variation is phonologically conditioned (e.g. 'compression' conditioned by the number of available syllables), or whether any of it is meaningful" (p. 145). Before GlatoBI is used for large scale transcription, it will be important to try to resolve these questions on the basis of closer study of the different patterns of transcriber disagreement.

3.7. Evaluation of study

The levels of agreement achieved with GlaToBI tended to be very slightly lower than those achieved in the ToBI and GToBI tests. We suggest two possible explanations for these differences. Unlike the evaluations of ToBI and GToBI which used read speech as well as spontaneous speech, the evaluation of GlaToBI used only spontaneous speech. Spontaneous speech generally suffers more from disfluencies (hesitations, false starts, etc.) and unclear (or completely flat) pitch contours than does read speech. As a result, the intonation of spontaneous speech is more difficult to transcribe. Another possible explanation for the slightly lower levels of agreement is the fact that the scale of the current evaluation is much smaller than those carried out for both the original ToBI system and the GToBI system: ToBI: 489 words, 26 transcribers; GToBI: 733 words, 13 transcribers; GlaToBI: 273 words, 7 transcribers. The effect of one or two transcribers disagreeing with the others would therefore be greater in the GlaToBI evaluation than in the two earlier studies. Taking all of this into account, the levels of agreement reached by transcribers using GlaToBI are in fact quite comparable to those seen in both the original and the German ToBI systems.

REFERENCES

[1] A. H. Anderson, M. Bader, E. G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. S. Thompson, and Weinert R. The HCRC Map Task Corpus. *Language and Speech*, 34(4):351–366, 1991.

[2] M. E. Beckman and G. M.. Ayers. *Guidelines for ToBI Labelling, Version 1.5*. Ohio State University, 1993.

[3] Alan. Cruttenden. Rises in English. In J. Windsor Lewis, editor, *Studies in General and English Phonetics: Essays in Honour of Professor J. D. O'Connor*, pages 155–173. Routledge, 1994.

[4] M. Grice, M. Reyelt, R. Benzmüller, J. Mayer, and A. Batliner. Consistency in transcription and labelling of German intonation with GToBI. In *Proceedings of the 1996 International Conference on Spoken Language Processing*, 1996.

[5] D. R. Ladd. *Intonational phonology*. Cambridge University Press, 1996.

[6] J. Pierrehumbert. *The Phonology and Phonetics of English intonation.* PhD thesis, MIT, 1980.

[7] J. Pitrelli, M. Beckman, and J. Hirschberg. Evaluation of prosodic transcription labeling reliability in the ToBI framework. In *Proceedings of the 1994 International Conference on Spoken Language Processing*, volume 1, pages 123–126, 1994.

[8] K. Silverman, M. Beckman, J. Pitrelli, M. Ostendorf, C. Wightman, P. Price, J. Pierrehumbert, and J. Hirschberg. TOBI: A standard for labelling English prosody. In *Proceedings of the 1992 International Conference on Spoken Language Processing*, 1992.