Investigating the Syntactic Characteristics of English Tone Units

Alex Chengyu Fang, Jill House, and Mark Huckvale
alex,jill,mark@phon.ucl.ac.uk

Department of Phonetics and Linguistics, University College London
Gower Street WC1E 6BT, London, England

ABSTRACT

This paper describes an investigation into the correspondence between grammatical units and English tone units. Our first aim is to provide some statistics based on scripted read speech since past studies mainly dealt with spontaneous speech. The second aim is to investigate whether the clause structure is a reliable indication of the tone unit. We start with a description of the annotation of transcribed speech data selected from the Spoken English Corpus (SEC), which is tagged for detailed wordclass information with *AUTASYS* and then parsed for rich syntactic description with the Survey Parser. Prosodic annotations in SEC, including both major and minor tone unit boundaries, were then mapped onto the parse trees. We then present our observations of tone units in the light of the clause structure. The paper will demonstrate that there is an overall correspondence between the clause structure and the tone unit in the sense that tone units generally co-start with the clause and that they seldom occur at major clause element junctures.

1. BACKGROUND

While acknowledging that tone units are "readily adjustible to the demands of emphasis, grammatical complexity, speed of utterance, and other factors" [1:1359], past studies have established an overall correspondence between tone units and grammatical units [2, 3, 4]. This correspondence has been extensively exploited in text-to-speech synthesis systems for the segmentation of written texts into tone units (TU). Altenberg [5:306], for instance, reported a performance of 85% predictive accuracy for a set of grammatical rules. However, rule-based prediction tends to draw on sophisticated syntactic analyses for clause elements such as subject, predicate, object, and adverbial, which are not always available from the generally problematic automatic parsing.

Svartvik and Fang [6] describe the design and implementation of a software system, *SpeechMaker*, which automatically chunks English text into tone units. *SpeechMaker* is also a rule-based system that depends on grammatical and syntactic information for its TU prediction. Internally, it integrates wordclass analysis and syntactic parsing for appropriate segmentation of the text into TUs. In general, the system implements the idea that there is correspondence between TUs and grammatical units in the sense that TUs mainly occur between clause boundaries and rarely interrupt the internal structure of the canonical grammatical phrases (adjective, adverb, noun, preposition, and verb). In particular, the overall design was inspired by the understanding that "the verb and its noun phrases constitute not only a syntactic unit but a semantic unit, one referred to as a *proposition*" [7:11]. In the light of this, one may view the tone unit as a phonological manifestation of the proposition itself, which is syntactically expressed through, for instance, combinations of clause elements such as the subject-predicator-complement-adverbial (SPCA) construction depending on the verb valency. From this, it follows that boundaries of constructions such as SPA, SPC, and SPCA may be interpreted as TU boundaries. It also follows that verb valency alone may be used to determine such boundaries without a global parse. Indeed, *SpeechMaker* is instructed by only five rules for its TU prediction:

- Finite transitive and copular VPs must be grouped together with the antecedent NP and the complement.
- Finite intransitive VPs must be grouped together with the antecedent NP and the following PP or AVP.
- Sentence-initial prepositional phrases (PP) and adverb phrases (AVP) are by default co-extensive with a prosodic unit boundary.
- Relative pronouns, relative adverbs, and conjunctions (coordinating and subordinating alike) never occur at the end of a unit.
- Heavy NPs, defined as equal to or longer than half of the maximum TU length, are assigned a default prosodic unit boundary. In *SpeechMaker*, the maximum TU length is set to nine words and a heavy NP is in effect one with four or more words.

To apply the rules above, *SpeechMaker* analyses the input text for detailed wordclass and sub-categorisation information which serves as a basis for subsequent parsing at the phrase level. The verb phrase, which inherits its transitivity from the main verb, then determines the clause boundary and sets it as a potential TU boundary before the length factor is applied to finalise the prediction.

These rules have all been implemented in *SpeechMaker* and have yielded satisfactory results with various test texts. However, their validity had not been empirically checked against corpus data. Results from past studies of the correlation between syntax and prosody may be conveniently borrowed to justify the use of these rules but none of these studies reported results from read speech. The main purpose of our study, therefore, is to conduct a similar investigation based on read speech and find out to what extent the clause structure as a reflex of the proposition corresponds to the TU. In the following

sections, we shall first of all describe the corpus and its annotation before results are presented and discussed.

2. CORPUS DATA

2.1. Text Selection

The corpus selected and used in our study is the Spoken English Corpus (SEC), compiled at the University of Lancaster in conjunction with the Speech Research Group at the IBM UK Scientific Centre [8]. The corpus contains orthographic transcriptions from British radio broadcasts, which are annotated for various prosodic properties such as major and minor tone unit boundaries. The first three texts of the commentary category were selected: `SECA01`, `SECA02`, and `SECA03`, plus the second text from magazine-style reporting, `SECF02`. The punctuated version of these texts was tagged, parsed, and then manually corrected before the prosodic version was used to map tone unit (both minor and major) boundary indications onto the parse trees.

2.2. Wordclass Tagging

An automatic wordclass tagger, *AUTASYS* [9], was used to tag the punctuated version of the selected texts with a tagset designed for and used by the International Corpus of English (ICE, [10]). A major reason for the use of the ICE tagset is that it provides detailed information regarding lexical sub-categorisation. It notes 16 general word classes, all of which, except for four, are described by a feature set that indicates their sub-categorisations and grammatical and/or morphological forms. There are a total of 270 possible class-feature combinations. The verb class, for instance, is tagged for seven different transitivity types as well as six different grammatical forms:

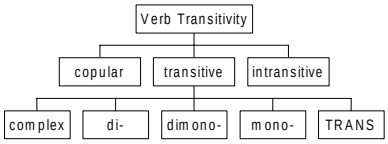

Figure 1: ICE verb transitivity scheme, with transitive verbs subdivided into five further types.

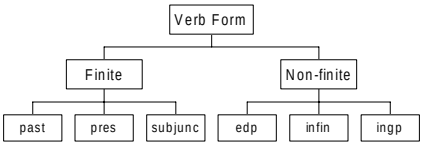

Figure 2: ICE verb form

The `TRANS` type refers to verbs complemented by non-finite clauses with an overt subject. Its various constructions are illustrated by Examples (1)-(4).

(1) I <u>asked</u> her to go.
(2) They <u>had</u> the room painted.
(3) He <u>made</u> me feel nervous.
(4) John <u>saw</u> himself spending a fortune on that useless thing.

2.3. Syntactic Parsing

The tagged texts were then parsed for syntactic information with the Survey Parser [11, 12]. Again, the parsing scheme is that used for ICE, which not only specifies grammatical categories such as the clause, AJP, AVP, NP, PP, and VP but also their syntactic functions (also known as grammatical relations, such as subject, verb, and object). Table 1 is a summary of the major syntactic functions and their realisations.

Functions	***Sub-Functions***	***Realisations***
Subject	subject	clause, NP
	provisional subject	NP (anticipatory *it*)
	notional subject	clause
Verb		VP
Object	direct object	clause, NP
	indirect object	NP
	provisional object	NP (anticipatory *it*)
	notional object	clause
Complement	subject complement	AJP, AVP, clause, NP, PP
	object complement	AJP, AVP, NP, PP
	transitive complement	clause
	focus complement	clause
Adverbial		AVP, clause, NP, PP

Table 1: A summary of the major clause elements in the ICE parsing scheme.

Minor entities mainly include those belonging to the internal structures of the five phrase types. The phrase-internal structures are shown in Figure 3.

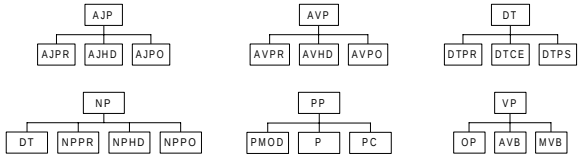

Figure 3: The ICE phrase-internal structures.

The verb phrase normally inherits its transitivity feature from the main verb (`MVB`). When it is passivised, however, the verb phrase undergoes the following valency changes that affect the clause structure:

complex transitive	→	copular	(Example 5)
dimono-transitive	→	intransitive	
di-transitive	→	mono-transitive	(Example 6)
mono-transitive	→	intransitive	(Example 7)
TRANS	→	mono-transitive	(Example 8)

Examples (5)-(8) illustrate these changes.

(5) His language <u>is called</u> Yathoyua and there's a dialect of it that I speak or that I write anyway <u>called</u> Dhamyathua.
(6) You would do if you <u>were given</u> the opportunity.
(7) The opportunities didn't seem to be part of the way I <u>was brought</u> up and <u>educated</u>.
(8) We <u>were recommended</u> to do this at university by one of the lecturers.

2.4. Tone Unit Mapping

Finally, TU boundaries in the prosodic version of the texts were mapped onto their corresponding syntactic trees. The final version of the parse trees is illustrated by Figure 4:

Figure 4: A syntactic tree augmented with TU boundaries.

3. ANALYSES

3.1 Basic Statistics

There are 147 sentences in the corpus with 2,716 words, which yields an average sentence length of 18.5 words. A total of 629 tone unit markers (with minor and major TUs collapsed) were observed, resulting in 4.3 tone units per sentence and 4.3 words per tone unit. These basic statistics are summarised in Table 2.

Text	***S***	***W***	***TU***	***W/S***	***TU/S***	***W/TU***
SECA01	41	767	155	18.7	3.8	4.9
SECA02	40	731	202	18.3	5.1	3.6
SECA03	33	606	140	18.4	4.2	4.3
SECF02	33	612	132	18.5	4.0	4.6
Total	147	2716	629	18.5	4.3	4.3

Table 2: Sentence and TU lengths in words.

Table 3 lists statistics regarding the frequency of occurrence of the major clause elements. Statistics about the clause are also listed for discussion in the following sections. The third row lists frequencies of tone units that occur before such elements. The complement (*C*) column glosses together objects and complements in Table 1. Note that frequency counts for the clause overlap with those for complements and adverbials (c.f. Table 1).

	S	***P***	***C***	***A***	***Clause***
Frequency	279	355	250	322	386
TU Start	117	95	40	109	289
%	41.9	26.8	16.0	33.9	74.9

Table 3: TU breaks before the major clause elements

3.2 Clauses and TU Boundaries

Results indicate that clause boundaries do in general correspond to TU boundaries. Firstly, all of the 147 sentences in the corpus are set off from each other by a TU marker. Secondly, of all the 386 clauses in the corpus, 289 are fronted by a TU marker, a percentage of 74.9%. This seems to confirm, in frequential terms, the observation that the clause being co-extensive with the tone unit "can be regarded as the unmarked or neutral state of affairs" (Halliday 1966:120). Where this norm does not apply, some definable patterns can be identified. For instance, clauses not directly fronted by TUs tend to be either verbal or prepositional complements, incorporated into a larger grammatical structure. Indeed, when this happens, the TU beginning tends to fall directly in front of the verb or the preposition. According to our data, 60.8% of the remaining clauses may be accounted for this way: of the 97 remaining clauses not fronted by TUs, 47 are used as verbal complements (see Examples 9 and 10) and 12 as prepositional complements (see Example 11). In these circumstances, the grammatical cohesion forces the tone unit beginning to fall elsewhere.

(9) Some Free Church people | feel <u>that in practice | the Anglicans go it alone whenever they can</u> ‖

(10) At one point | the forms asked <u>us to state | our last contact | with the organisation</u> ‖

(11) It's the notion of <u>an impersonal | uncaring universe | relentlessly following laws of cause and effect</u> ‖

It is also worth noting that TUs co-starting with the beginning of a clause account for almost half (45.9%) of the total number of tone units in our data, thus making clause the singular category that correlates with the largest number of tone unit starts among all the grammatical entities concerned in the study.

3.3 Clause Elements and TU Boundaries

Another indication of the clause as co-extensive with TUs is the small number of major clause elements separated from each other with TU markers. In particular, the predicator is rarely separated from its subject and its complement by the beginning of a tone unit. In our data, there are altogether 355 predicators, of which only 95 or 26.8% are fronted and thus separated from their subjects by TU markers. More than half of these cases (56 in all), moreover, may be accounted for by the following factors:

- **Antecedent apposition**: In announcing the award in New York | the rector of the university | <u>Dr Nicholas Argentato</u> | described Mr Moon as | a prophet of our time ‖
- **Heavy subject**: Next week | <u>a delegation of nine Protestant ministers | from Argentina</u> | visits the Autumn assembly | of the British Council of Churches ‖
- **Heavy complement**: Some Free Church people | <u>feel that in practice | the Anglicans go it alone whenever they can</u> ‖
- **Emphasis**: But the scientists <u>themselves</u> | weren't having any of that ‖

- **Intruding adverbial or parenthetical clause**: The Christian view they argue | is to see events that can be covered by natural laws | as God's usual activity ||

Discounting the above, the subject-predicator interruptions by TU markers are reduced to 35, or only 9.8% of the total number of predicators in our data.

An even smaller number of predicates are separated from their complements by a TU marker, i.e., only 40 or 11.2%. The rarity of such cases is expected since the predicate normally forms a strong syntactic cohesion with the complement. Indeed, with the 106 NPs as direct objects, only 9 (8.5%) are fronted with a TU marker. With clauses as complements, however, there is a tendency for them to co-start with a tone unit. According to our data, nearly 50% (23 out of 49) of complementing clauses are observed to co-start with tone units.

Jointly, breaks occurring at SPC junctures account for only 15.3% of the cases, slightly lower than Crystal's 19 per cent [3:262]. They represent 21.8% of all the TU markers observed in the corpus. It is significant that a majority of such cases may be accounted for by the length factor.

4. CONCLUSION

Results from our investigation seem to confirm, in frequential terms, the correspondence between tone units and the clause in the sense that a majority of clauses co-start with tone units. While it is not clear whether most tone units also terminate at the end of clauses, our empirical observation has nevertheless shown that tone units seldom interrupt major clause element junctures such as subject, predicator, and complement, which may be interpreted to indicate that most tone units do terminate at least after the obligatory elements of a clause. Within the clause, there is a varying degree of cohesion between the elements. The predicator, for instance, seems to form a stronger link with its complement than with its subject.

Though based on scripted read speech, our results regarding the major clause elements are surprisingly similar to those presented by Crystal [3], whose work was mainly based on transcribed spontaneous speech. This similarity between spontaneous and read speech indicates that the organisation of tone units is mainly centred on the propositional structure as the basic information unit syntactically manifested through the clause and is therefore unaffected by the varying degrees of syntactic complexities. In the light of our own findings, our observations deviate from Crystal [3] regarding the correlation between the clause and the tone unit: the frequency evidence in our data clearly indicates that a large majority of clauses co-start with tone units. This argument can be further strengthened by the fact that tone units seldom start at major clause element junctures. A most forceful indication will be results from a future experiment designed to investigate whether the end points of tone units also correlate to the closure of clause structures.

From a synthesis point of view, the results confirm the validity of the rules designed for *SpeechMaker*. The use of verb valency to determine the clause boundary as an indication of potential TU breaks provides a simple but reliable approach towards automatic TU segmentation. The length factor, on the other hand, may be effectively used to account for exceptions to this general rule.

5. ACKNOWLEDGEMENT

The work was supported in part by the Engineering and Physical Science Research Council, UK, Grant Nos GR/L52109 and GR/L81406.

6. REFERENCES

1. Quirk, R., S. Greenbaum, G. Leech and J. Svartvik. *A comprehensive grammar of the English language*, Longman, London, 1985.

2. Quirk, R. J. Svartvik, A.P. Duckworth, J.P.L. Rusiecki and A.J.T. Colin. "Studies in the correspondence of prosodic to grammatical features in English," *Proceedings of the Ninth International Congress of Linguistics*, 679-691, 1964.

3. Crystal, D. *Prosodic systems and intonation in English*, Cambridge University Press, Cambridge, 1969.

4. Altenberg, B. *Prosodic Patterns in Spoken English*, Lund University Press, Lund, 1987.

5. Altenberg, B. "Predicting Text Segmentation into Tone Units," J. Svartvik, editor, *The London-Lund Corpus of Spoken English – Description and Research*, 275-286, Lund University Press, Lund, 1990.

6. Svartvik, J. and A.C. Fang. "SpeechMaker," T. Nevalainen and L. Kahlas-Tarkka, editors, *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*, 431-450, Société Néophilologique, Helsinki, 1997.

7. Jacobs, R.A. *English Syntax: A Grammar for English Language Professionals*, Oxford University Press, Oxford, 1995.

8. Knowles, G., B. Williams, and L. Taylor. *A Corpus of Formal British English Speech*, Longman, London and New York, 1996.

9. Fang, A.C. "AUTASYS: Automatic Tagging and Cross-Tagset Mapping," [10], 110-124.

10. Greenbaum, S. (Editor). *Comparing English World Wide: The International Corpus of English*, Oxford University Press, Oxford, 1996.

11. Fang, A.C. Automatically Generalising a Wide-Coverage Formal Grammar. In C. Percy, C. Meyer, and I. Lancashire, editors, *Synchronic Corpus Linguistics*, 131-146. Rodopi, Amsterdam, 1996.

12. Fang, A.C. "The Survey Parser: Design and Development," [10], 142-160.