Glossa a journal of general linguistics

Simard, Candide and Claudia Wegener. 2017. Fronted NPs in a verb-initial language – clause-internal or external? Prosodic cues to the rescue!. *Glossa: a journal of general linguistics* 2(1): 51. 1–32, DOI: https://doi.org/10.5334/gjgl.211

RESEARCH

Fronted NPs in a verb-initial language – clause-internal or external? Prosodic cues to the rescue!

Candide Simard[1] and Claudia Wegener[2]

[1] SOAS, University of London, Russell Sq., London WC1H 0XG, UK

[2] University of Cologne, 50923 Köln, DE

Corresponding author: Candide Simard (cs75@soas.ac.uk)

This paper investigates prosodic features of fronted constituents in the verb-initial Oceanic language Gela (spoken by about 16.000 people in Solomon Islands). Although Gela's basic constituent order is verb-(object-)subject/predicate-subject, constituents can appear in front of the verbal predicate. Fronted constituents in Gela can be interpreted as pre-clausal (i.e. external to the following clause, immediately preceding it) or clause-initial (i.e. clause-internal, at the very beginning of the clause), each of which can be associated with certain information structure categories of topics and focus. This paper discusses how prosody provides clues towards the interpretation of fronted constituents as pre-clausal or clause-initial, based on a quantitative study of their prosodic correlates. We argue for using prosodic criteria established on clear examples to help analyse ambiguous cases. The results are compatible with an approach that recognises the importance of prosody in syntactic analysis and contribute data from a little known language to the discussion to what degree prosodic and syntactic phrasing are aligned.

Keywords: prosody; information structure; word order; Oceanic languages; prosodic phrasing; syntactic phrasing

1 Introduction

This paper investigates prosodic features of fronted constituents in the verb-initial language Gela (spoken by about 16.000 people, one of about 70 Oceanic languages in Solomon Islands). The basic constituent order in Gela is VOA/VS/PS (Miller 1974: 470; Wegener in prep.), as shown in example (1).

(1) dp_clips_nlg_101

[(E)	tutu-a]$_{VC}$	[na	ngali]$_{O}$	[na	vaivine]$_{A}$.
3SG.PST	pound-3SG.O	ART	ngali.nut	ART	woman

(Speaker EG: What is she doing now?) 'The woman is pounding the Ngali nuts.'

However, constituents can appear in front of the verbal predicate. This conforms to the pragmatic discourse tendency of fronting certain constituents, either by topicalisation or focusing that has been observed cross-linguistically in verb-initial languages (Payne 1999). Fronting is a common feature in the area, mentioned for example in the grammatical descriptions of Lengo (Unger 2008), Longgu (Hill 1992; 2002), and Toqabaqita (Lichtenberk 2008), but the prosodic features of different types of topical and focal constituents that can be fronted have not been thoroughly investigated yet.

We identify two pre-verbal positions in Gela: a clause-initial position (i.e. clause-internal, at the very beginning of the clause) and a pre-clausal position (i.e. external to the following clause, immediately preceding it). The pre-clausal position (PC) is used for shifted/contrastive left-dislocated topics or for “free” or “hanging” topics (Maslova & Bernini 2006). Constituents in this position do not fill any syntactic position in the core clause, in the sense that they are not arguments or adjuncts; they are at most coreferent with an argument or adjunct, see Section 3.1.1 for details. As for the clause-initial position (IP), constituents in this position always fulfil some syntactic function within the clause and are either arguments in focus or contrastive topics in one particular construction used to express parallel contrast, see Section 3.1.2 for details.

Looking at natural data, the analysis is clear and unambiguous for only a small number of examples, e.g. cases where both the PC and IP positions are filled. If there is only one fronted constituent, the analysis can still be straightforward, e.g. when the fronted constituent is in focus (then it must be inside the clause, i.e. in IP position), or when it is a left-dislocated topic coreferential with a pronoun inside the clause (then the topic is outside the clause, i.e. in PC position, as Gela does not permit two coreferent NPs in different positions of the same clause, e.g. before and after the verb). Section 3.1 below provides an overview of the different functions of fronted NPs in Gela, and discusses in more detail how different morphosyntactic and information-structural criteria can be used to analyse any given example. We show that in many cases these criteria are not sufficient to achieve an unambiguous analysis because the overt expression of arguments is not obligatory in Gela, and also because the identification of information structure categories is notoriously difficult in natural data (Cook & Bildhauer 2011). This has led us to investigate the prosodic features of the constituents in these two positions with the aim of assessing whether they display distinct prosodic patterns. In such a case, these prosodic patterns could help in the analysis of morphosyntactically ambiguous cases, and could, in all likelihood, be used by hearers to interpret utterances correctly. This investigation is interesting not only from a practical point of view, but also for theoretical considerations as it contributes data from a little known language to the ongoing research aiming to decide whether (or how) different patterns in prosodic phrasing correspond to different levels of syntactic phrasing. It further brings forth the question whether prosodic features could, or should, have a place in the definition of morphosyntactic constructions.

Consistent with existing research on prosodic phrasing and boundary strength (Gee & Grosjean 1983; Lehiste 1983; Wagner 2005; Xu 2009), it is assumed that a prosodic boundary between adjacent constituents tends to be stronger when the syntactic boundary between these constituents is stronger, in the sense that they reflect different levels of syntactic dependency. Prosodic boundaries are defined here in reference to prosodic information in the signal. Boundary strength relates categorical differences between prosodic boundaries of units of different size to the gradient differences in the acoustic cues encoding prosodic juncture. We hypothesize that, if there is one at all, a prosodic phrase boundary between the syntactic constituents within one clause (i.e. between the verbal predicate and the rest of the clause, or between the IP position and the verbal predicate) should be weaker than a prosodic phrase boundary at a syntactic clause boundary (i.e. between the PC position and the clause). Additionally, we hypothesize that the prosodic phrasing of a focused IP constituent is similar to that of a clause-initial verbal predicate because both are in focus and both are constituents within the clause.

Evidence is provided in the form of an instrumental analysis of the prosodic cues for the encodings of IP and PC. An empirical investigation of the prosodic properties of these two syntactic positions is presented, based on datasets extracted from narratives, procedural texts and utterances elicited by means of visual stimuli.

This paper is organised as follows: Section 2 is a review of prosodic phrasing in linguistic theory; it also introduces the prosodic theory used here. Section 3 provides a presentation of Gela, an overview of the functions of the two pre-verbal positions, and information on the data analysed. We also sketch the methodology used in this study. In Section 4, the results of the research are presented. Section 5 discusses these results, and Section 6 concludes the paper.

2 Prosodic phrasing

The role of prosody as a basis for the organisation of content in speech, distributed on the basis of prosodic constituency, is generally acknowledged. Speech consists of utterances which comprise words grouped together into phrases. These in turn form increasingly larger units, resulting in syntactic and prosodic units of different sizes. These units are signalled, in the acoustic domain, by particular intonation, accent or timing patterns, which use cues such as pauses, changes in the amplitude of intensity and pitch, and lengthening of the final syllables to mark boundaries.

Linguistic theories deal differently with prosodic constituency. In formal theories the interface between prosody and syntax has been formalised in different ways (e.g. Klein 2000, HPSG; Szendrői 2003, Minimalism; Bögel et al. 2009, LFG). In the Chomskyan tradition, for instance in its most recent formulation, the Minimalist Program (Chomsky 1995; 2000), only two levels of representation are assumed: Phonological Form (PF) and Logical Form (LF). Sentences are generated by the computational system which takes the numeration (a set of lexical elements) as input and creates a syntactic structure (D-Structure, S-Structure). This structure is linearized and assigned a prosodic structure at PF, and is then interpreted at LF. However, because of the non-isomorphy of prosodic groupings and syntactic constituents (Zec & Inkelas 1990), and because of the close relation between discourse features and prosody, the level where prosody should be represented is still widely discussed (Fanselow 2007; Slioussar 2007; Elfner 2012, *inter alia*). Two different kinds of non-isomorphism raise problems for this kind of "direct syntax" approach: non-isomorphism between prosodic and syntactic constituents, and non-isomorphism between syntactic structures and intonational meaning.

Functionalist accounts, on the other hand, point out form-function correlations, generally aiming to motivate or explain form through function. In these approaches, the study of prosody is integrated with the study of grammar and meaning in natural social interactions. One of its aims is to describe particular languages in such a way that speaker behaviour can be predicted (phenomenological description). In a framework such as Construction Grammar (CxG), for instance, it is assumed that phrasal constructions are learned pairings of form and meaning (Goldberg 1995; 2006). Specific constructions are licensed by phonological factors that include intonation.

The aim of this paper is not to prove or disprove a certain formal theoretical view on the syntax-prosody interface; our aim is to present a description of our data from a lesser known language and the patterns we find at the syntax-prosody interface. In our opinion, all linguistic theories need data and descriptions for as many diverse languages as possible to be able to test hypotheses based on each theory's predictions, and to improve their theoretical framework. Our work thus aims to be of use to any theory interested in the phenomena discussed in this paper.

The definition of prosody supported here has both functional and formal aspects. We make use of the Parallel Encoding and Target Approximation model (PENTA, Xu 2005) which assumes that arbitrary language specific rules exist in prosody which are dependent

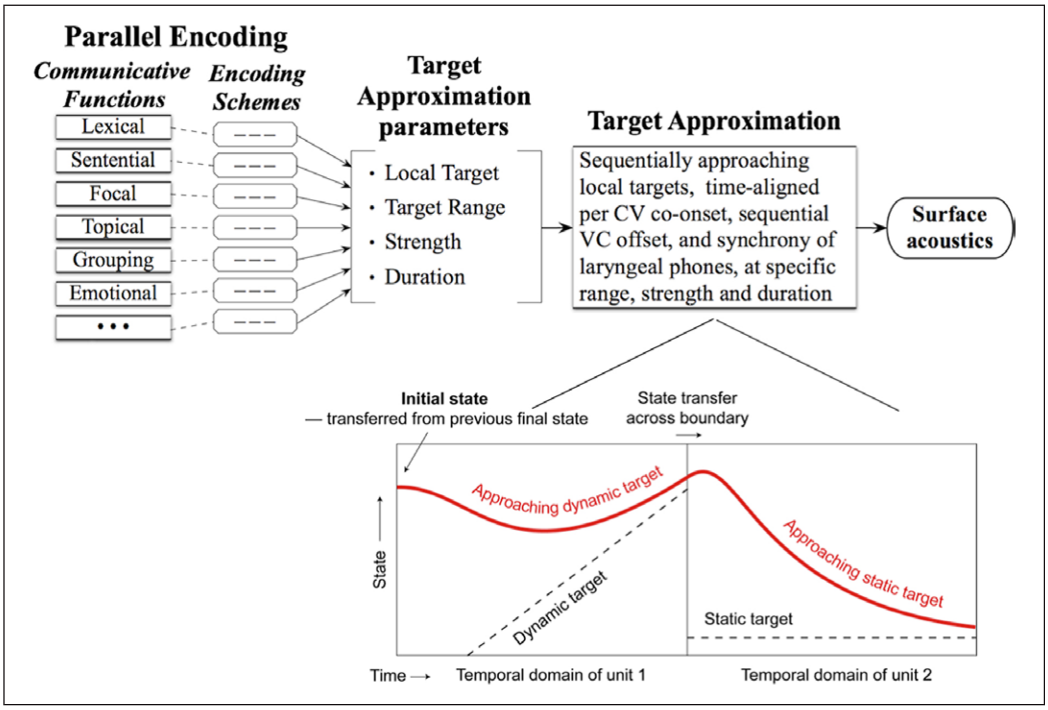

Figure 1: Upper panel: A schematic sketch of the PENTA model. Lower panel: The target approximation component of PENTA, which is an articulation process (from Xu et al. 2015).

on functions, for instance stress pattern of words, focus or modality. The model is presented schematically in Figure 1 (from Xu et al. 2015). It takes as its basis communicative functions (on the left) which are transmitted in parallel through encoding schemes that are the result of the interaction of all the communicative functions involved – note that the inventory of communicative functions need be empirically determined for each language (Xu 2005). The encoding schemes are transmitted via a limited number of parameters that are viewed as phonetic primitives. Four primitives are recognized for speech melody: local pitch targets, pitch range, articulatory strength and duration (middle box in Figure 1). This process generates surface acoustics, including F0, through the mechanism of target approximation, TA (lower panel in Figure 1).

TA is the mechanism by which pitch targets are assigned to each syllable. They are conceived as the underlying pitch trajectory resulting from their articulatory approximation by speakers. That is to say, a speaker produces F0 which approaches the target from the onset and throughout the syllable, but because of the limits of the articulatory process (Xu 2005) the target may not be fully realised before the speaker starts on the next syllable. On the graph, syllable boundaries are indicated by vertical lines; the underlying pitch targets are shown with a dashed line, and the pitch is represented with the thick curve.

Pitch targets can be static, for instance [high], [mid], and [low], or dynamic [rise] and [fall]. Target slope, height, duration and strength are the parameters proposed to capture quantitatively the characteristics of pitch targets (Prom-on et al. 2009). The slope and height specify the form of the pitch target; for example, the Mandarin rising and falling tones are found to have positive and negative slope values, respectively (Prom-on et al. 2011). Height is calculated relative to the speaker F0 mean; strength indicates how rapidly a pitch target is approached: the higher the strength value, the faster F0 approaches the target.

Finally, we make use of the notion of intonation unit (IU) and follow Chafe (1987) in defining it as a stretch of speech uttered under a single intonation contour, delimited by pauses, changes in tempo, and other prosodic cues, and usually (but not always) corresponding to a clause.

3 Gela

Gela belongs to the Oceanic subgroup of the Austronesian language family, more precisely to the Guadalcanal-Gelic subgroup of the Southeast Solomonic languages. Most of its 16,000 speakers (Solomon Islands National Statistical Office 2009) live in the Florida Islands, but one district of Savo Island (about 30 km away) has been home to a few hundred Gela speakers for some generations. The language spoken in the other districts of Savo Island is Savosavo, an unrelated Papuan language. Especially on the Florida Islands, Gela is the main language of everyday interaction, it is still transmitted to children, and there are still monolingual speakers. The main threat to its vitality comes from Solomon Islands' Pijin (English-based pidgin creole/extended pidgin), which is used more and more by the younger people. The varieties of Gela spoken in both locations are currently being documented as part of Wegener's research on language contact between Savosavo and Gela. The total corpus collected as part of this research consists of about 25 hours of transcribed Gela recordings, with interlinear glossing for about 12 hours. A subset of recordings (total length 1h 52min) was annotated for the purposes of this study; for more detail on this data see Section 3.2 below.

The syllable structure in Gela is C(V). Word stress generally occurs on each word's penultimate syllable, but the addition of morphemes after the root may cause the penultimate syllable to become unstressed (Crowley 2002). Preliminary studies have shown that a stressed syllable has higher F0 and intensity, and longer duration than its unstressed counterpart. According to Miller (1974: 470) and a few other short existing descriptions, basic constituent order in Gela is VOA (see Crowley 2002 and references therein); this was confirmed by our own data (see Wegener in prep.). Other syntactic features associated with verb-initial languages such as prepositions and articles before noun order are also observed. Gela exhibits low degrees of synthesis and fusion, i.e. it is mildly agglutinating and has very few portmanteau morphemes. It has a number distinction between dual and plural and frequently makes use of serial verb constructions (SVC, two or more verbs which together form one verbal predicate and share tense and aspect marking); both of these features are common in the area and shared by many Austronesian as well as non-Austronesian languages (cf. e.g. Foley 1986; Lynch, Ross, & Crowley 2002). Syntactic subjects and objects are cross-referenced in the verb complex, subjects by means of a proclitic portmanteau tense and subject marker that marks the beginning of the verb complex, and objects by an object suffix attaching to the verb. The subject proclitic can occasionally be dropped, especially at the beginning of a clause, or in procedural texts without a specific subject referent. Syntactic function is not indicated on additional subject- and object-NPs, which are in fact often dropped. Example (1) above is one of the few instances of a clause with overtly realized subject and object NPs. The subject proclitic *e* '3SG.PST' was not produced in this case, but is added in round brackets to illustrate the usual structure.

This research is part of the project "Discourse and Prosody across Language Family Boundaries", which investigates information structure categories and prosody in Gela for the first time. Until recently, while there has been some work on Polynesian languages (cf. e.g. Bauer 1991; Calhoun 2015) these aspects of grammar have received scant attention in the descriptions of Oceanic languages in Melanesia. Notable exceptions are Clemens (2014), who argues that sentential constituents can be reordered to satisfy constraints on prosodic well-formedness in Niuean, an Austronesian language of the Tongic subgroup, two chapters on topicalisation and focus in the grammar of Toqabaqita, a Southeast Solomonic language from the Longgu/Malaita/Makira subgroup (Lichtenberk 2008), and a study of the prosodic marking of focus in Torau (Jepson 2014).

3.1 Pre-verbal positions, functions of constituents

While Gela's basic constituent order is VOA/VS, this paper focuses on those constituents that can be fronted, either in a pre-verbal position at the beginning of a core clause which can be filled by any argument, referred to as clause-initial (IP); or in sentence-initial position that precedes the core clause, similar to what is found in a great number of other languages (Van Valin 1993: 6). It is referred to as pre-clausal (PC), that is, it is external to the following clause, immediately preceding it. NPs in this location do not fulfil any syntactic function in the core clause, but can be co-referent with one of the arguments or adjuncts.

In example (2) both the PC and the IP position are filled.[1] It is the answer to a content question 'What does the young man take?' The topic, *the young man*, is expressed in PC position, with a co-referent pronoun later in the core clause that functions as the subject. The object NP *the shark* is in clause-initial IP position, followed by the verb complex. In this and all following examples topic constituents will be bold, and focus constituents in small caps.

(2) sj_pictures_nlg_107

PC			[IP		[PRED]$_{\text{Pred}}$		X]$_{\text{C}}$
Gari	**mane**	**ke**,	[NA	BAGHEA	[te	hola-a]$_{\text{Pred}}$	**gaia**]$_{\text{C}}$.
child	man	EMPH	ART	shark	3SG.NFUT	take-3SG.O	3SG

(EG: What does the young man take?) '**The young man**, **he** takes THE SHARK.'

If both positions are filled, as in this example, the analysis of the sentence is straightforward. Sections 3.1.1 and 3.1.2 provide an overview of the morphosyntactic and information structural features of these two positions, illustrating them with the help of clear and unambiguous examples. Section 3.1.3 then shows examples where the syntactic analysis is not straightforward.

3.1.1 Functions of constituents associated with the PC position

The constituents in PC position are either left-dislocated topics or hanging topics. The NP *gari mane* 'young man' in example (2) is a left-dislocated topic: the participant is expressed first and then resumed in the main clause using a pronoun. Left-dislocation is associated with shifting topics, implying a contrast between this and other active topics, as noted in Foley & Van Valin (1985: 356).

The PC position with the same information structural function is also found with non-verbal clauses, as in example (3). The nominal predicate is in focus, and the IP position is not filled. The analysis of the first NP as being in PC position is clear because the subject of the clause is the pronoun *gaia* 'it' that is coreferent with the left-dislocated NP.

The intonation contour of (3) (Figure 2) shows how the first constituent forms a unit bounded by a pause and ending with a pitch peak, followed by a second unit set off by a pitch reset and displaying the clausal contour associated with declarative sentences in Gela (rising from the beginning to the end of the predicate and falling over the rest of the clause).

[1] In this and all following examples, the boundaries of the core clause and the predicate (verb complex or predicate NP) are indicated by square brackets with the respective subscripts *C* and *Pred* respectively if they can be clearly identified.

(3) ls_ti_mt_3_008
A **eni** **ke**, [[NA BULUKA]$_{\text{Pred}}$ **gaia** **ke**]$_{\text{C}}$.
ART 3SG.PROX EMPH ART COW 3SG EMPH
(Man & Tree task, starting to describe the next picture)
This one, **it** (is) THE COW.’

The other possible function of an NP in PC position is that of a hanging topic. In this case “[…]the topic expression is juxtaposed to a clause-like component denoting the main proposition, and does not specify a variable of this proposition […]” (Maslova & Bernini 2006: 10). In example (4), the core clause consists only of an intransitive verb complex with a serial verb construction and the subject NP *na mane* ‘the man’. The initial NP *sakai* ‘one’ refers to the picture which is described using this proposition.

The intonation contour (Figure 3) shows once again the initial IU ending with a pitch peak and a pause, and the pitch reset and normal clausal contour of the core clause in the second IU. Note that long pauses are found to occur not only in elicitations based on pictorial stimuli, but also in narratives, which leads us to consider them as part of Gela’s prosody and not solely artefacts of elicitation tasks.

(4) fs_lo_mt_AS_138
Sakai ke, [[E RIU SAPA]$_{\text{Pred}}$ **na** **mane**]$_{\text{C}}$.
one EMPH 3SG.PST turn move.seawards ART man
(Man & Tree task, starting to describe the next picture)
‘As for one (picture), **the man** FACES SEAWARDS.’

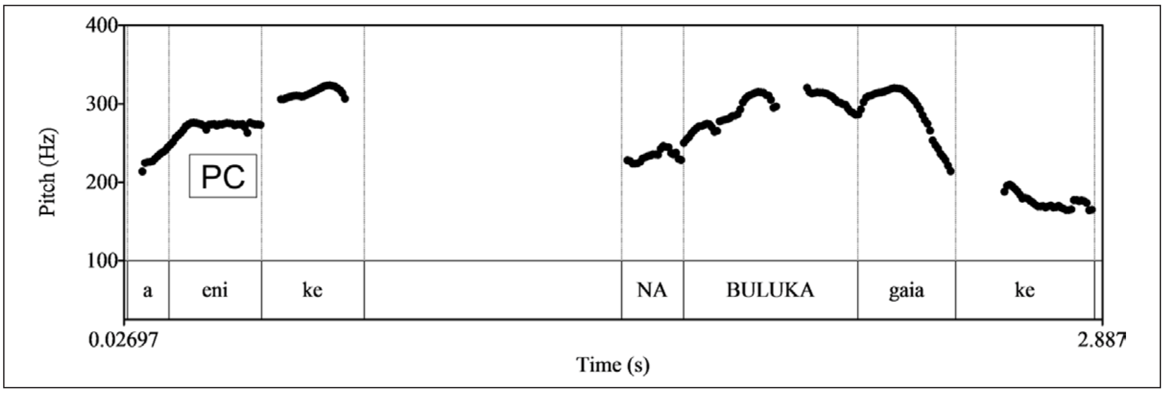

Figure 2: Intonation contour of example (3).

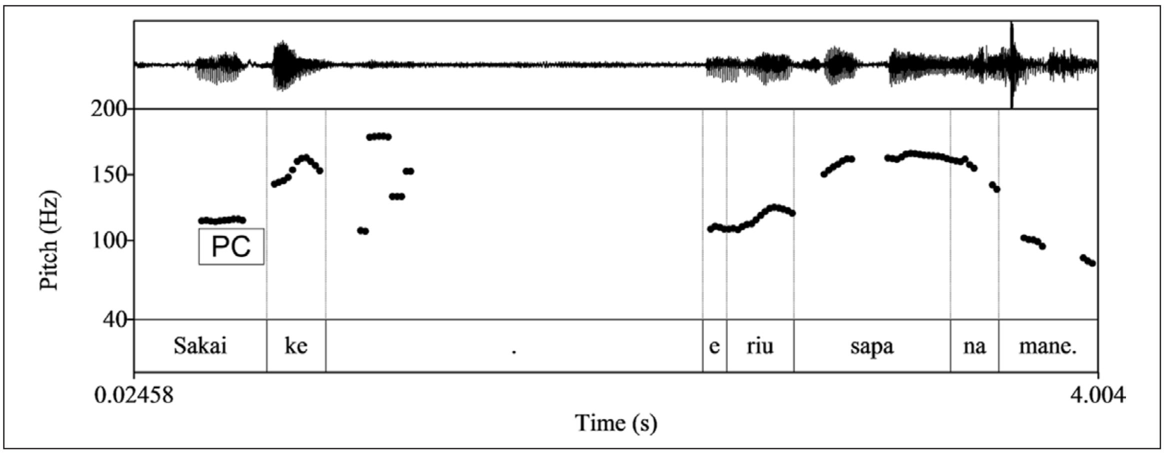

Figure 3: Intonation contour of example (4).

The analysis here is straightforward because the referent of the initial NP is not a participant of the subsequent clause. It could not even be interpreted as a locative adverbial, as in this case it would have to be encoded as a prepositional phrase with the locative preposition *ta* 'LOC'.

Based on the examples so far, one might get the impression that the emphatic enclitic marker =*ke* is an indicator of a topic in PC position. However, example (3) above shows that it is not restricted to this position, but can also appear on a clause-final topic expression. In addition, targeted elicitation and discussions with speakers showed that it is not obligatory in any position. The fact that it is very often found on NPs in PC position can therefore only be used as circumstantial evidence supporting a PC-position analysis, but not as a sufficient criterion.

3.1.2 Functions of constituents associated with the IP position

An NP in IP position can be either argument focus (Lambrecht 1994; narrow focus in Van Valin & LaPolla 1997), a subtype of information focus where one argument is the only new information in the sentence, or a contrastive topic in a parallel contrast construction.

The domain of argument focus is usually limited to a single NP constituent. Operationally, it can be identified as a target to a wh-question. The function of the argument focus is to provide the missing argument in a presupposed open proposition, for example the NP *the shark* in example (2), repeated here for convenience.

(2) sj_pictures_nlg_107
Gari mane ke, [NA BAGHEA [te hola-a]$_{Pred}$ **gaia**]$_{C}$.
child man EMPH ART shark 3SG.NFUT take-3SG.O 3SG
(EG: What does the young man take?)
'**The young man**, **he** takes THE SHARK.'

IPs also appear in constructions in which two clauses with a parallel structure are used to give contrasting new information for two contrasting topics. Repp (2016) notes that "[…] contrastive topics are often viewed as topics with a focus (e.g. Büring 1997; Krifka 2007), where focus is viewed as the information-structural category that elicits alternatives to the focused element […]" (see also Rooth 1985; 1992). Example (5) shows this construction, in which the speaker describes a picture with three bananas. After saying that there are three bananas, he produces two coordinated clauses, first to explain the orientation of two of the bananas (which are both facing towards the east), and then the orientation of the third banana (which is facing seawards instead). The intonation contour (Figure 4) shows two IUs corresponding to the two clauses, in the first, the pitch rises throughout the IP, reaches a plateau and then falls sharply at the end; in the second IU, the contour is much less pronounced.

(5) bp_ad_mt_3_116
[**E rua** [TORO RIU HORU]$_{Pred}$]$_{C}$
ART two DU.NFUT turn descend
m- [**e sakai** [TE RIU SAPA]$_{Pred}$ (...)]$_{C}$.
and- ART one 3SG.NFUT turn move.seawards
(Man & Tree task, describing the positioning of three bananas)
'(Three bananas.) **Two** face downwards (towards the east) and **one** faces seawards […].'

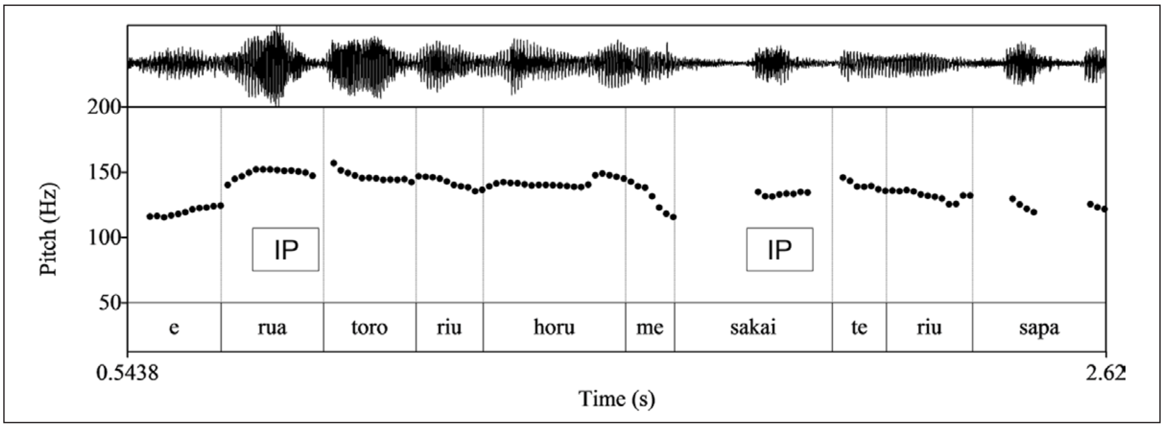

Figure 4: Intonation contour of example (5).

In this construction it is not possible to have any coreferent material after the predicate, which clearly shows that the pre-verbal constituent is part of the core clause.

Note that the parallel construction is characterized by both topic and focus being contrasted, as in ***John***$_{\text{ContrTop}}$ *bought* PEANUTS$_{\text{ContrFoc}}$, ***Peter***$_{\text{ContrTop}}$ *bought* BANANAS$_{\text{ContrFoc}}$. The shifted topics discussed above in 3.1.1 are also contrastive in some sense, as they imply other topics available for discussion, e.g. other characters in a story – in this case, however, there is no one corresponding contrasting focus tightly linked with each topic. Staying with the example of characters in a story, there are several new bits of information one could add for each character, without any two of them being in direct contrast. According to our analysis of the Gela data to date, we observe that contrastive topics are only expressed within the core clause if they are paired with a contrastive focus, thus forming part of the parallel construction.[2] If only the topic is contrastive, it is provided in PC position and can then be picked up as a normal topic by a clause-final coreferential pronoun within the core clause (cf. examples (2) and (3) above). Given that the parallel construction described above is such a marked context, it is relatively easy to identify as long as both clauses are produced, making examples such as (5) easy to identify and analyse.

3.1.3 Ambiguous cases

We have seen that the two pre-verbal positions can have different functions:

- PC Left-dislocated (shifted/contrastive) topic
  Hanging topic
- IP Argument (narrow) focus
  Contrastive topic in parallel contrast construction

Leaving aside the well-known problem of identifying information structure categories in naturally occurring speech (Cook & Bildhauer 2011), for the PC position, morphosyntactic criteria in combination with information structure render examples with a hanging topic unambiguous: the NP in PC position has no syntactic function in the clause, the verb complex does not show agreement with it, nor is it marked as an adverbial prepositional adjunct. It simply provides the backdrop for the upcoming proposition. For the IP position, we also have one information structural function that allows for a clear analysis, the function of argument focus. If the pre-verbal NP provides the only

[2] But note this does not hold in the opposite direction, i.e. it is possible to have a contrastive focus without a contrastive topic, e.g. when for a given topic different actions or objects are contrasted.

new information in the clause, it has to be part of the clause, and thus can only be in IP position.

However, topics that are contrastive to some degree are found in both positions. Usually they would be placed in the PC position, allowing for coreferential material within the following clause (e.g. *Peter, he* SLEPT), but this material is often not there (e.g. *Peter,* (*he*) SLEPT). The parallel contrast construction is the only context where a contrastive topic is found in IP position, but apart from the striking structural parallelism of the two clauses of this construction (*John* ATE; *Peter* SLEPT), there are no morphosyntactic markers within the clause indicating this construction. These possibilities are the main source for ambiguous examples, particularly in natural speech and in more or less uncontrolled contexts. Note that this ambiguity is unlikely to cause an unsurmountable obstacle for a native speaker, whether the topic is weakly contrastive and the following focus a basic information focus, or the topic as well as the focus are contrastive, in either case the first constituent is given and topical, and the second new and in focus. Given a normal communicative situation, the hearer will probably be able to tell from the context which construction (and thus interpretation) fits better; and even if not, the difference in meaning is so subtle (on the level of presuppositions, implications and implicatures) that most interactions can simply continue and the ambiguity be resolved through the following utterance(s).

Examples (6) and (7) are illustrations of clauses for which it is hard to decide whether the initial NP should be analysed as part of the clause. Example (6) shows an ambiguous verbal clause. There is no formal indication as to whether it should be interpreted as a PC '**(As for) the young child**, **(he)** is picking out the ngali nuts […]', or as an IP '**The young child (in contrast to the older child)** is picking out the ngali nuts (in contrast to cracking the nuts open) […]'

(6) dp_clips_nlg_089

Gaia na gari pile ke [E VILI-RA]$_{Pred}$
3SG ART child small EMPH 3SG.PST choose-3PL.O
NA NGALI…
ART ngali.nuts
(EG: What is the small child doing?)
'**The young child**, (**he**) IS PICKING OUT THE NGALI NUTS […]'
Or
'**The young child** IS PICKING OUT THE NGALI NUTS […]'

Similarly, example (7) shows an ambiguous non-verbal clause which could be interpreted either as '**(As for) me**, (**I** am) a snake'; or '**I (in contrast to you)** (am) a snake (in contrast to a human).'

(7) rr_cs_likuliku_023

Inau ke [NA POLI]$_{Pred}$.
1SG EMPH ART snake
(Snake woman speaking: 'Oh darling, me, I am not a human being, that you would come and call me!')
'**Me**, (**I** am) A SNAKE'.
Or
'**I** (am) A SNAKE.'

3.2 Data and methodology

Datasets were extracted from the existing partial documentation (narratives and recordings of the *Man & Tree* games,[3] the latter have the code *mt* in the file name), supplemented by data elicited by asking questions about visual stimuli (i.e. pictures of people and things, represented by the code *picture*, cooking video clips, represented by the code *clips*) designed in order to elicit comparable data on rarer information structure categories (such as argument focus and parallel contrast constructions). This elicited material still consists of spontaneous utterances, inasmuch as the answers were not scripted; speakers were free to answer the questions as they wanted. The recordings are stored in the DoBeS online archive,[4] and the coded transcripts will be added in due course. Recordings were made using a Zoom H2 digital audio recorder and a Beyerdynamic MCE82 stereo condenser microphone.

As mentioned earlier, one difficulty with spontaneous speech data is that there is always a certain degree of uncertainty in the analysis of the information structure of a given clause, even in the elicited conditions. False starts, repairs and noise in a token can be identified and lead to a token being excluded from analysis, but structural ambiguity was harder to deal with. The selection of tokens was done by the second author, who conducted the data collection and has some fluency in the language, and verified by the first. A set of Verb Initial clauses (VI, predicate focus) was selected to serve as benchmark on which to base our comparison: they were chosen as "default" constructions because they represent a neutral topic-comment structure without particular emphasis on either topic or comment.[5] Therefore the VI contours were used to compare the datasets of representative tokens for PC and IP positions. A sample of ambiguous tokens was also selected (coded as IP-PC_unsure).

Out of a total of 135 coded examples, altogether 76 tokens were selected from 6 male speakers; this somewhat limited number of tokens is due to the difficulty of selecting examples from natural speech that can reasonably be interpreted as truly comparable. Those 76 tokens comprise 8 that have constituents in both PC and IP positions (coded as PC_IP). As shown in Table 1, clear examples of PC are few (5), but to this number must be added the 8 tokens that have both a PC and an IP, for a total of 13 tokens. The two types of constructions were coded differently, in case of significant differences (which

[3] This task was originally developed to elicit spatial expressions in a game setting (Pederson et al. 1998). Two speakers sit side by side, separated by something so that they cannot see each other, with identical sets of 12 or 16 photos spread out in front of them. One speaker is the director, the other is the matcher. The director describes each photo in his set in whatever order she likes, and the matcher has to identify each picture in his own set. Several of the photos show the same objects in slightly different configurations and positions, encouraging a detailed description of position and orientation of the objects to be able to distinguish them.

[4] The Gela part of the archive can be found here: http://hdl.handle.net/1839/00-0000-0000-000D-B68F-C@view

[5] The possible interaction between the placement of phrase boundaries and the length of the constituent need be considered, as pointed out by an anonymous reviewer. It has been noted that speakers tend to chunk speech into phrases of comparable size, making prosodic phrasing sensitive to what has been called "balancing demands" (Fodor 2002). Sensitivity of this kind has been reported in a variety of languages (D'Imperio et al. 2005). However, languages vary in how they balance "prosodic weight" (number of syllables) and syntactic complexity. For example Elordieta et al. (2003) found that Catalan presents a tendency to divide utterances into phrases of similar syllabic lengths, but Spanish shows a strong tendency to separate S from the rest of the utterance material. Jung (2003) found, for Korean, that syntactic phrases made of up to five syllables form a phrase, but stretches of six or more syllables form two (see also Elfner 2012 for Irish). The VI in our datasets may be less syntactically complex than the PC/IP examples, being composed of either VS, VO or VAdv/VObl clauses. They are similar to the PC/IP in that there is usually only one phrase after the VC, but they may be less long because they obviously have at least one less phrase in front of the VC. However, given that phrases in our Gela examples vary greatly in the number of syllables they contain, there is no evidence that phrasing is motivated by "prosodic weight", in the Catalan style.

Table 1: Number of tokens of each subtype in the datasets.

Clause Internal (IP)	20
Pre-Clausal (PC)	5
Unsure	20
Verb Initial (VI)	23
With 2 constituents: Pre-Clausal and Clause-Internal PC_IP_both)	8
Total	76

Table 2: Information structure categories associated with subtypes.

		Type		
		IP (incl PC_IP)	PC (incl PC_IP)	VI
Information Structure Categories	TOPabout	1	7	0
	TOPhang	0	6	0
	TOPcont	5	0	0
	FOC_PRED	0	0	23
	FOC_ARG	19	0	0
	FOC_CONTR	3	0	0
	Total	28	13	23

were not found). There are 20 IPs to which must be added the 8 PC_IP tokens, totalling 28 tokens. A number of ambiguous tokens (20) were selected as well; finally, 23 VI tokens complete the datasets. The distribution of tokens in each subtype is equally shared among the speakers, except for the VIs which were extracted from 4 instead of 6 speakers. Preliminary ANOVA tests showed no significant effect for a category "speakers" on the measurements.

Table 2 shows the Information Structure categories associated with the tested subtypes. Topics can be "aboutness" (TOPabout, this code was used for left-dislocated topics), "hanging" (TOPhang), or "contrastive" (TOPcont, these are the strongly contrastive topics in a parallel contrast construction). The scope of the focus can be on the argument (FOC_ARG) or the whole predicate (FOC_PRED), or focus can be "contrastive" (FOC_CONTR) (in opposition to the argument and predicate focus which are "information" focus).

As noted before, the VIs are all instances of predicate focus. IPs are more likely to be associated with argument focus (19), contrastive focus (3) or contrastive topics (5). PC are either aboutness, i.e. left-dislocated (7), or hanging (6) topics.

Field-based data collection always raises challenges, and the quality of sound recordings vary in our datasets; most contain inevitable background noise which impacts on prosodic analysis. But despite the difficulties and even if it is challenging, we are convinced that, while it is easier to work with controlled scripted data, it is still possible to conduct prosodic analysis based on less controlled data: patterns should be discernible in natural speech, otherwise speakers and hearers would not use them. As an illustration, Figure 5 shows the mean F0[6] in the Verb Initial tokens. The values on the left of the vertical line in the graph correspond to the first constituent, here the verb as predicate focus, showing

[6] A note of caution about means: they may not be robust against outliers.

syllables at the boundaries: first, second, penultimate (penult) and final syllables. At the right of the vertical line are the same syllables in the constituent that follows the verb, usually an NP. A clear pattern emerges at first glance, even if individual tokens diverge from it.

This pattern is clearly evidenced in example (8), Figure 6, showing the same rise until the end of the first constituent in *e riu longa,* no pitch reset and the falling contour in the second constituent, here *na mane.*

(8) fs_lo_mt_AS_262

[[E	riu	longa]$_{Pred}$	na	mane]$_{C}$.
3SG.PST	turn	move.landwards	ART	man

'The man is facing landwards.'

The selected datasets described above were analysed on the basis of acoustic criteria which are considered to be evidence of prosodic cohesion; they may refer to local or global prosodic events and can co-occur in a prosodic group (Ladd 1996), they include: presence of a final pause, coherence in global trends of F0 and energy (e. g. declination of both F0 and energy), reset at the beginning of a new phrase, or lengthening of final

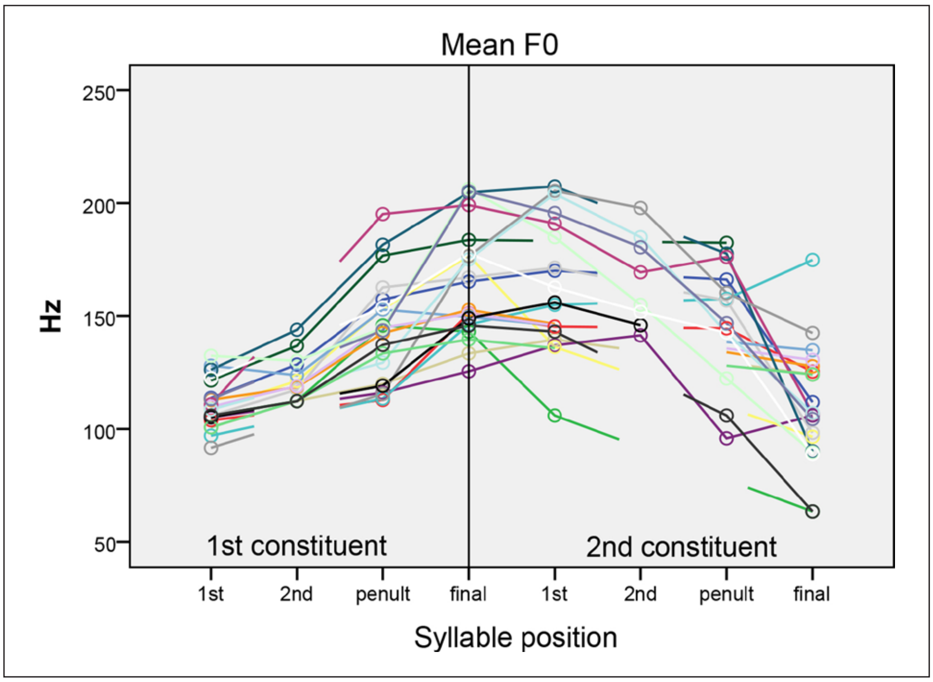

Figure 5: Mean F0 of the first, second, penultimate and final syllables in the first and second constituents of the tokens in the Verb Initial dataset. Each line represents a token.

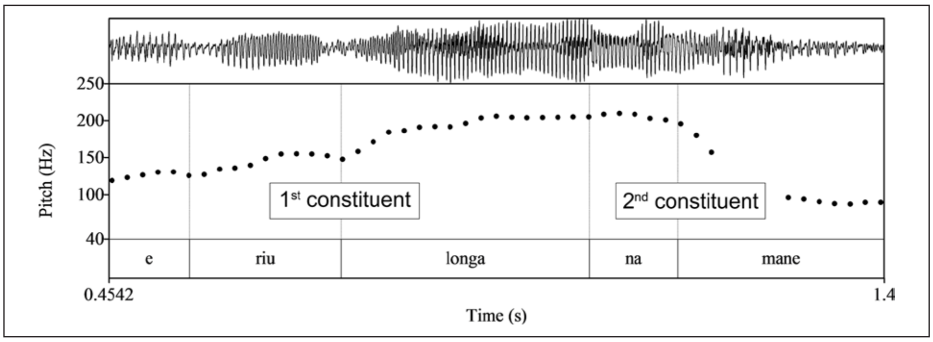

Figure 6: Intonation contour of example (8).

syllables. The analysis was based initially on the simultaneous inspection of waveforms, spectrograms, and pitch contours using PRAAT (Boersma & Weenink 2013), followed by the instrumental analysis, in order to facilitate the examination of the correlates at fine-detailed level, and for performing systematic statistical comparisons. All selected tokens were initially segmented into syllables; errors in vocal pulse markings were manually corrected. Speech rate and segmental content were not controlled; although it is well established that an increase in speaking rate correlates with a decrease in pauses and more generally on prosodic boundary marking (Goldman-Eilser 1968), we consider that speech rate is under speaker control in a way that absolute pitch is not, and can be purposely varied to achieve communicative aims. For the purpose of this analysis, only tokens evaluated on a perceptual basis as being uttered with a "normal" rate were selected, and no normalization was applied. The PRAAT scripts ProsodyPro (Xu & Prom-on 2014), and PENTAtrainer 1 (Prom-on, Xu & Thipakorn 2009) were then used to obtain the measurements of the prosodic correlates that include duration (in milliseconds), intensity (in decibels), mean F0 (average of 10 measurements over the syllable, in Hz), and excursion size (F0maxima – minima, in semitones).

In PENTA, several parameters are used to describe the F0 trajectory of each syllable, all of which are useful to specify the form of the pitch target[7]: height (final F0, in semitones), slope (final velocity, in semitones/second, and strength of target approximation (second^{-1}). Final velocity measurements are taken near the end of a syllable because PENTA assumes that this is where the pitch target is best approximated; semitones, rather than Hz, are used for pitch correlates (except for mean F0) as log scales have been shown to be more representative of auditory perception in linguistic contexts than Hertz (Nolan 2003), allowing us to generalize across speakers with different individual pitch ranges. Pauses – any gap in the waveform that is 30ms or more is labelled as a pause – whether they are present or not, and their length, were also taken into account, as well as final lengthening, considered a quasi language-universal phenomenon (Lehiste 1983; Wightman et al. 1992; Wagner 2005; Turk & Shattuck-Hufnagel 2007; *inter alia*). It predicts that syllables that are close to a following prosodic boundary tend to have longer durations than in other positions, ceteris paribus. Although the details of the phenomenon are still discussed, it has been observed that deeper boundaries seem to trigger more lengthening of the preceding syllables than weaker boundaries do.

The measurements and encodings for the VI, PC, and IP were compared and the results validated with a statistical analysis. The patterns emerging for the clear cases were then used to analyse those tokens identified as ambiguous.

Table 3: The distribution and duration of pauses by subtype.

Location of pause	Duration Mean (ms)	N	Std. Deviation
after PC	1594.24	3	592.13
after PC in PC_IP	1845.96	4	419.82
after IP	407.08	4	174.08
after IP in PC_IP	288.55	1	.
after 1st const_unsure	906.34	7	779.43
Total		19	

[7] Some of these parameters have been developed in order to resynthesize the F0 trajectories in the QTa tool (Prom-on, Xu & Thipakorn 2009, Xu & Prom-On 2014).

4 Results

4.1 Pauses

Table 3 indicates whether or not pauses are found between the PC, IP, VI constituents and the rest of the IU, and their duration.

Note that the verbs in the VI datasets are never followed by a pause, but pauses are found both after IPs and PCs, albeit not in all tokens of either category. In those tokens that have both PCs and IPs, we coded whether the pause followed the PC or the IP. The duration of pauses that follow PCs are much longer at 1594.24ms to 407.08ms for IP. It is striking that the same pattern is observed when both positions are present in the tokens (PC_IP). The measurements of the duration for each are illustrated in boxplots in Figure 7. The difference between the subgroups is significant ($F\ (4, 14) = 4.080$, $p < .05$), notably between the PC and IP and "unsure". Pauses after PCs are especially long, but as shown in the standard deviation and in the boxplot (Figure 7), they also display a fairly large variation.

4.2 Correlates of the "default" VI

Sequences of phrases in Gela are usually integrated prosodically under a single intonation contour which shows its own internal structure. It displays a global rise and fall throughout the IU, the maximum F0 is reached at the end of the first constituent (here the verb) with subsequent prosodic words slightly downstepped to each other. The regularity of this pattern facilitates identification of utterance-internal phrase boundaries which often interrupt the declination pattern. The most common utterance-final pitch contour observed is a fall.

The graphs in Figure 8 show the results of the means of the measurements for the correlates of the first, second, penultimate and final syllables for the first and second constituents of the VIs, which are viewed as a "default" contour. A vertical line indicates the final syllable of the first constituent.

The measures of the mean F0 indicate an overall rise in the surface contour from the first syllables of the first constituent, here the verb (left of the vertical line), with a steeper movement in the penultimate than in the final syllables as shown by the excursion size

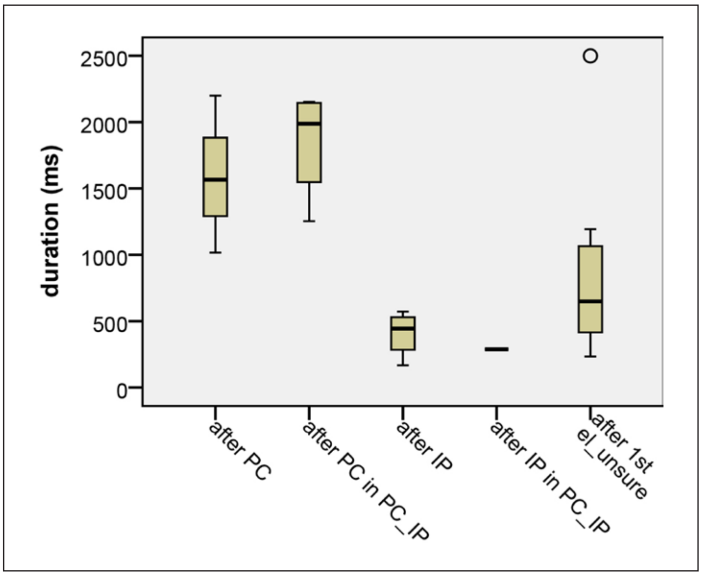

Figure 7: Boxplot showing the mean durations of pauses.

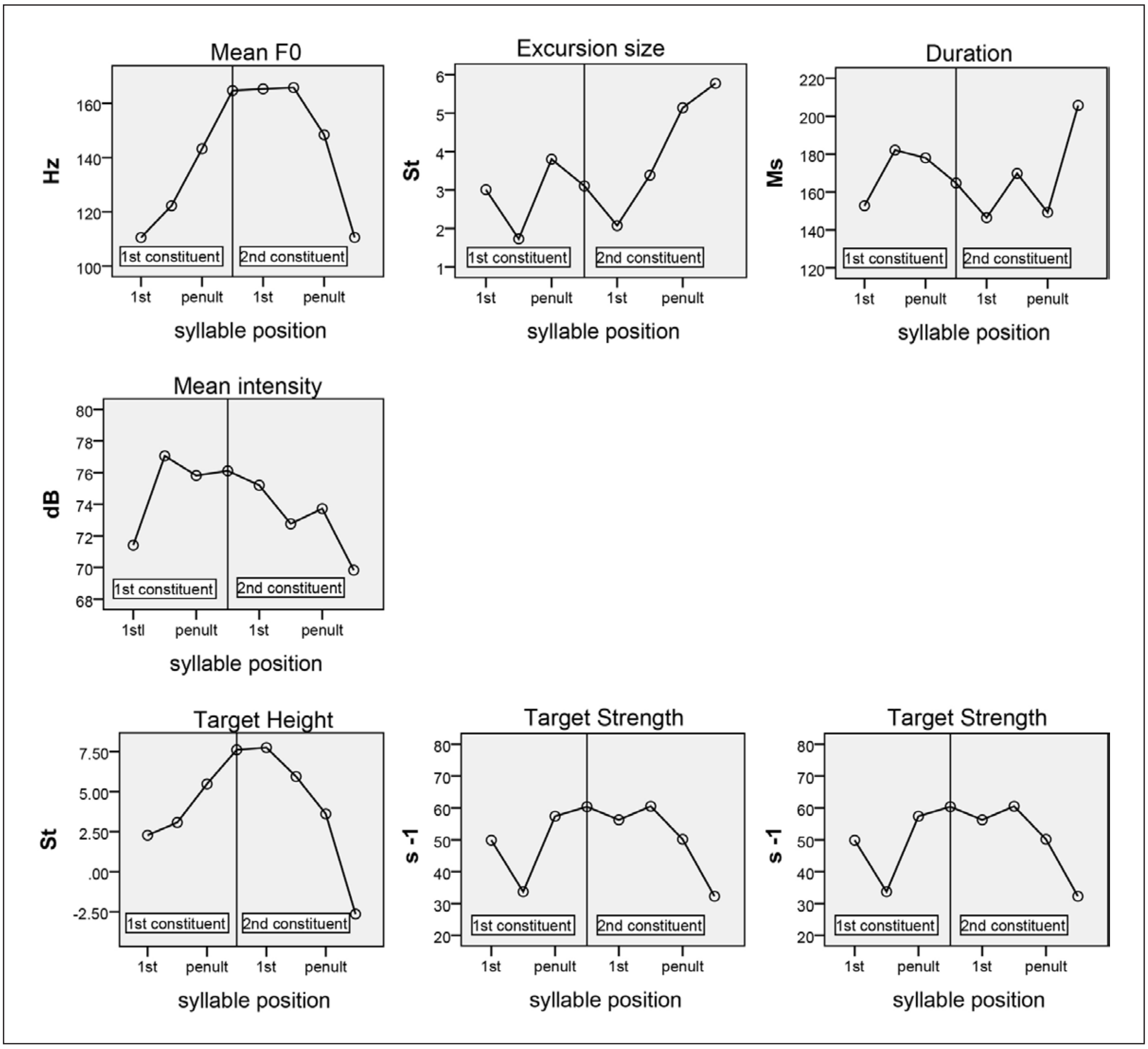

Figure 8: Prosodic correlates of VI dataset.

measure. The F0 movement over the end of the first constituent (verb) and the beginning of the following constituent correspond to the peak in the IU contour. Note that there is no pitch reset from the final syllable of the first constituent to the first syllable of the next constituent (after the vertical line). The bottom pane shows the values for the three parameters used in PENTA to define pitch targets. The measures for target height closely correspond to the absolute measures of F0. The graph of the target slope is interpreted in the following way: negative slopes occur due to F0 transitions from peaks towards the baseline, and positive slopes, from a lower, preceding position to F0 peaks. The second syllables in each of the constituents do not accord with the general patterns for the slopes measures: gently rising, 2.55st/s in the first syllable, with a rise to 3.91st/s in the penultimate syllable, followed by a fall to –9.73st/s in the final syllable of the first constituent, consistent with a prosodic word boundary, and then falling more sharply in the second constituent (–25.94st/s, –37.86st/s in the first and penultimate syllables) until finally reaching for the baseline in the IU final syllable (–1.3st/s). The measurements suggest salience at the right edge of the constituent in Gela: those of the target slope coupled with a greater pitch excursion and an increase in strength at which it is reached (target strength graph) highlight a sharp pitch movement in the penultimate syllable (corresponding to the stressed syllable). We posit a [high] pitch target in the final syllable of the first constituent, as a rising target would not have a negative slope.

As for the duration correlate, the final syllable of the first constituent is shorter than the penultimate, which suggests no final lengthening, which is expected at the end of phrases. The opposite is observed at the right edge of the second constituent where the final syllable is much longer than the penultimate. A similar intensity level is maintained from the penultimate to the final syllable of the first constituent; again, this is interpreted as an indication that the phrase boundary is not marked.

In short, the pitch, duration, and intensity patterns suggest no phrase boundary between first and second constituents in VI tokens. This is illustrated in Figure 9, where there is no evident discontinuity after the final syllable of *ghaghua* and the beginning of *gaia*.

(9) it_taulaghi_49
[[E ghaghua]$_{Pred}$ gaia na olomane]$_{C}$.
3SG.PST be.thus 3SG ART old.man
'Said the old man.'

4.3 A comparison of the measurements of IP, PC, and VI

Figure 10 shows the results of the measurements of the IPs and PCs, and VIs. For ease of reference, only the first constituents are shown, except in the graph of pitch reset, which includes the first syllable of the following constituent.

Considering the pitch measurements first: the results for the IPs and VIs resemble each other, particularly for the mean F0, the first, second, and penultimate syllables follow each other closely, rising steadily; the patterns differ in the final syllables where the mean F0 of the IPs declines slightly, but there is a sharp rise in the VIs. The pattern for the PCs is different; there is a gentle rise throughout, with a sharp rise in the final syllable. The measurement of the excursion size show that both PCs and IP have a greater excursion size at the right edge, but this is not evident in the VIs where the excursion size of the penultimate syllable is greater than that of the last syllable. The bottom graph of the top pane shows the measurements for the pitch reset, including the means of the first syllable of the second constituent: it shows a reset in both IPs and PCs (though more salient in PCs), but no reset in VIs. The excursion size and the pitch reset notably suggest phrasing in IPs and PCs, but not in VIs (see duration cue).

The measures of target slope, target strength, and target height suggest the following pitch targets at the right edge: a [high], or even a [mid] target for IPs, because of the weak strength (not a rise, because of the lack of increase in target height); and a [rise] target for PCs, suggested by the rise in target height, coupled with the sharp movement in slope and the increased strength.

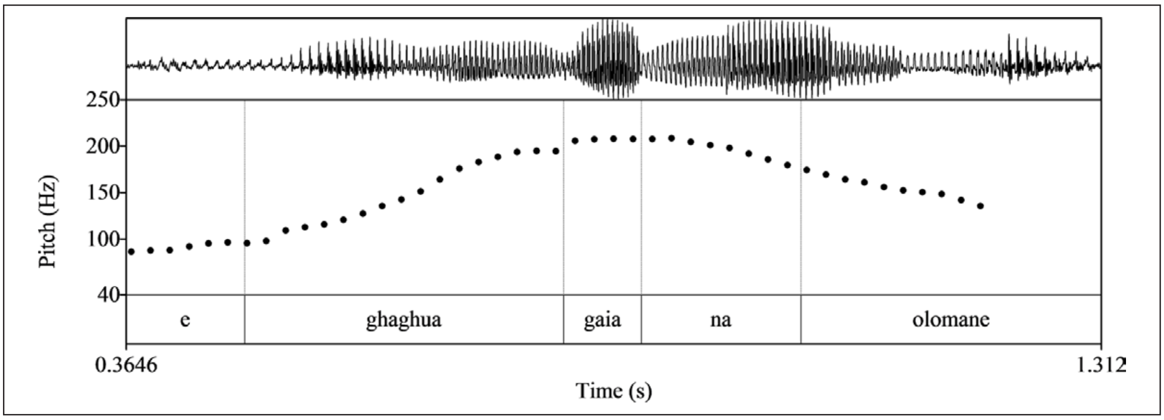

Figure 9: Intonation contour of example (9).

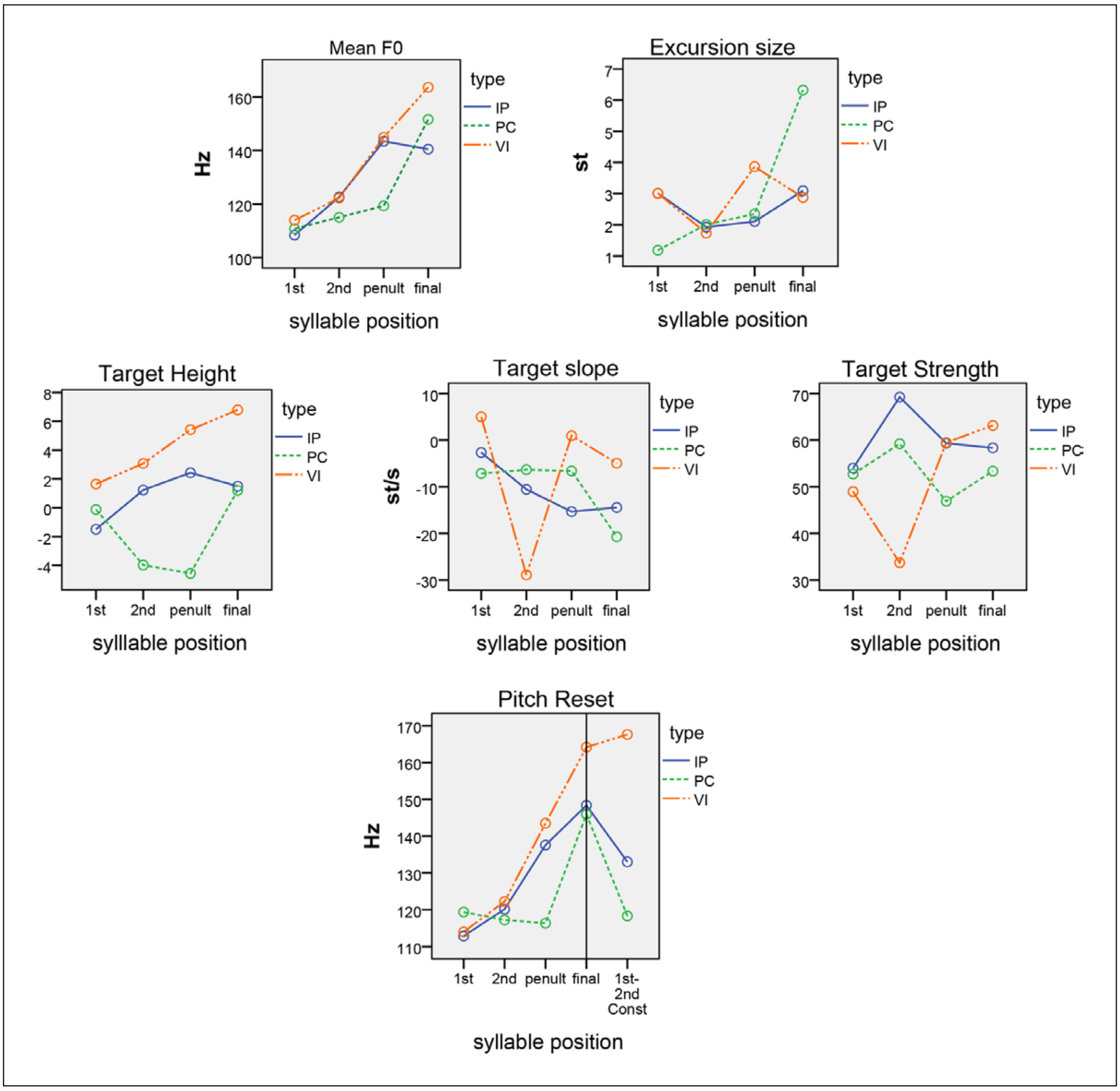

Figure 10: Comparison of the pitch correlates of the PC, IP, and VI.

Figure 11 shows the measures for the duration correlates. While the measures for VI and IP are somewhat close, the interesting patterns are found at the right edge. The duration of final syllables is longer than the penultimate syllables in IPs (178.08ms and 140.57ms) and PCs (279.35ms and 219.06ms) which suggests final lengthening, usually associated with the edge of a phrase; this is not apparent for the VIs which have averages of 174.19ms and 155.43ms, respectively. There was a statistically significant difference between the types as determined by one-way ANOVA ($F(2, 181) = 2.199$, $p = .037$). A Tukey post-hoc test revealed that the VI was significantly different from both IP and PC ($p < .001$). There were no statistically significant differences between the IP and PC ($p = .441$).

To formally assess whether the above observations of the graphical displays are significant, statistical tests were performed on the parameters (i.e., mean F0, excursion size, slope, height, and strength). A two-way repeated measures ANOVA was performed to test the effect of Type and Syllable Position on the measurements. Type has three levels (IP, PC, VI); Syllable Position has four levels (first, second, penultimate, final). The interaction between Type and Syllable Position was significant for all correlates, the means and standard deviations are presented in Table 4. The analysis of variance showed significant main effect for all the correlates for Syllable Position except target strength ($p > .05$) (Table 5).

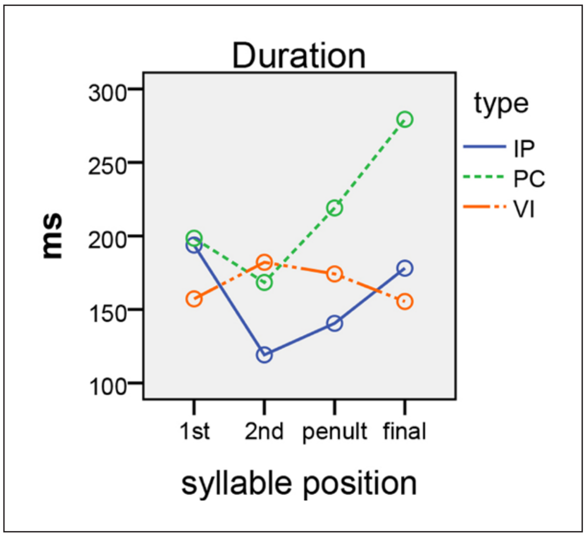

Figure 11: Comparison of the duration correlate of the PC, IP, and VI. Only the first constituents are shown.

Table 4: Statistical tests for the interaction between Type and Syllable.

Type x Syllable Position Dependent Variable	F	df	Standard deviation	Sig.
mean F0	1.309	2, 178	27.53	.035
excursion size	.608	2, 179	2.81	.05
target slope	2.225	2, 172	26.22	.008
target height	5.802	2, 181	7.67	.004
target strength	2.696	2, 181	22.29	.01
duration	2.199	2, 190	82.16	.037
intensity	3.715	2, 181	5.69	.026

Table 5: Statistical tests for Syllable Position.

Syllable position Dependent Variable	F	df	Sig.
mean F0	28.987	3, 178	.000
excursion size	5.296	3, 179	.002
target slope	3.325	3, 172	.041
target height	11.74	3, 181	.018
target strength	.654	3, 181	.581
duration	2.701	3, 190	.047
intensity	5.404	3, 181	.001

4.4 *Summary*

VIs have a higher mean F0 in their final syllables, a less sharp pitch movement, no pitch reset, and no final lengthening which suggests that their first constituent does not make a phrase of its own. On the other hand, the measurements for IPs and PCs suggest a phrase boundary at the right edge.

Comparing the IP and PC positions, it is found that the latter have a larger pitch reset than the former. Final syllables are longer in both IPs and PCs which indicate the edge of a phrase, but the lengthening is more extensive for the PCs. The measurements of target slope, height, and strength suggest different pitch targets at the right edge. A [high] for IPs because of the lack of increase in mean F0 and very minimal slope movement and reduced strength, which do not suggest a dynamic target; and a [rise] for PCs because of the sharp final F0, coupled with a large movement in slope and increased strength.

5 Discussion

Our findings indicate that the IPs, PCs, and VIs can be distinguished by a combination of parameters that include morphology, syntax, information structure, and prosody.

All three types of constituents, VIs, IPs, and PCs constitute an autonomous syntactic phrase. In Gela's canonical constituent order, the clause initial position is held by a verb, as in the VIs; the IPs are also placed in clause-initial position, but this position alone is not an adequate marker of discourse prominence; further prosodic marking is necessary as will be discussed below. Finally, the PCs are placed in utterance-initial position. On the morphological level, the PC (and rarely the IP) can be combined with the enclitic emphatic discourse marker =*ke*.

The constructions can also be distinguished by their information structural functions. The VIs are employed to express predicate focus, a constituent in IP position can be a narrow argument focus, or a contrastive topic in a parallel contrast construction, and a constituent in PC position can serve to express a shifted/contrastive left-dislocated topic or a hanging topic (i.e. it is a topic constituent that has no syntactic function in the following clause). This conforms to other analyses that have noted links between discourse pragmatics and clause structure, especially word order (see Clemens and Polinski to appear, for Austronesian and Mayan languages, and Givón 1983; Wagner 2005; Payne 1999; *inter alia*). For instance, it has been suggested that in some languages the default word order determines where given and new information is placed in a clause. Herring (1990) hypothesized that in verb-initial languages the topic (often the subject of the clause) would follow the verb and the focus would precede it. Thus, the "Word Order Principle" (Herring 1990: 164) states that information structure is determined relative to a language's basic word order as a rhetorical marking strategy. The link between clause constituent order changes and discourse structure has been reported for a number of languages, notably by Longacre (1995) in his study of the languages Trique of Mexico, Luwo in Sudan and Biblical Hebrew, and Palmer (2009) who showed that association of information structure categories with different positions in the clause lead to considerable word order variation in Cheke Holo, another Oceanic language of the Solomon Islands. Gela also conforms to this pattern.

The major findings reported in this paper concern the prosodic patterns associated with the IP and PC positions. As expected for our benchmark tester, the VI is fully integrated with the rest of the clause; the IP can constitute its own prosodic phrase, but its boundary is weak, and it is closely integrated into the intonation contour of the utterance it precedes. It can be followed by a pause, and typically displays a [high] pitch target at its right boundary. The PC also forms its own phrase, it is not closely integrated with the following IU; nonetheless its contour, as shown in the declination line, indicates that it forms a whole with the prosodic sentence. It is more likely to be followed by a pause than an IP, and pauses following a PC are longer than those that follow IPs. Pauses in Gela are long, nevertheless, the patterns observed conform to what has been reported for English, for example, in Cooper and Paccia-Cooper (1980), where fronting a constituent (their "preposed" category) can cause the following pauses to be lengthened by 50% to 200%. Finally, a PC typically displays a [rise] pitch target at its right boundary. These findings

Table 6: Summary of parameters used to identify the different constructions.

	VI	IP	PC
Syntax	autonomous phrase clause initial	autonomous phrase, rarely combined with emphatic discourse marker =*ke*	autonomous phrase, often combined with emphatic discourse marker =*ke*
		placed in clause-initial position	placed in an utterance-initial position
Prosody	fully integrated typically displays a [high] target at the right boundary	own IU, but not so strong boundary, integrated into the intonation contour of its clause	forms its own IU, but integrated in prosodic sentence
		can form a phrase with following subject agreement proclitic (misalignment of syntactic and prosodic phrasing)	more likely to be followed by a (longer) pause
		may be followed by a pause	typically displays a [rise] target at the right boundary
		typically displays a [high] target at the right boundary	
Function	predicate focus	narrow argument focus	shifted/contrastive left-dislocated topic
		contrastive topic in parallel contrast construction	hanging topic

are summarised in Table 6. The examples (3), (4) and (5) and their accompanying pitch tracks are illustrative of the prosodic characteristics described here.

We started with two hypotheses; firstly, we hypothesized that a prosodic phrase boundary between the syntactic constituents of one clause should be weaker than a prosodic phrase boundary at a syntactic clause boundary, and this was confirmed by our findings. The boundaries between both IPs and VIs and the rest of the clause are weaker than those between the boundaries between the PCs and the rest of the clause, reflecting their syntactic integration.

Additionally, we hypothesized that the prosodic phrasing of a focus IP constituent would be similar to that of a clause-initial verbal predicate, because both are in focus and both are constituents within the clause. However, our findings have disproved this hypothesis. We found that the encodings for the VIs indicate a phrase that is much more closely integrated with the rest of the clause than in the case for the IPs. However, further testing is needed in order to fully disprove our hypothesis, firstly we need to test separately the IPs that serve as arguments in focus and those that serve as contrastive topics; secondly, it would also be interesting to distinguish between different types of VIs, those where the predicate focus actually only contains the verb complex, i.e. intransitive clauses, and those where the predicate focus also extends over a following object NP.

5.1 Testing with "unsure" cases

In a final step, the correlates discovered for the clear cases can be used to test those cases labelled as "unsure" because none of the morphosyntactic and information-structural criteria can unambiguously identify the initial constituent as in IP or PC position.

For this purpose, a *k*-means cluster analysis is used. Cluster analyses such as *k*-means explore data in statistical ways so that items grouped into a cluster are more similar to one another than to items in different clusters. The measurements of the prosodic correlates were used as input for the *k*-means cluster analysis. The findings of the analysis are then

Table 7: Test 1: 3 clusters, correlation of cluster and type, counts and percentages.

Cluster		Type			Total
		IP	PC	VI	
1	Count	1	4	0	5
	%	3.6%	30.8%	0.0%	7.8%
2	Count	24	3	21	48
	%	85.7%	23.1%	91.3%	75.0%
3	Count	3	6	2	11
	%	10.7%	46.2%	8.7%	17.2%
Total	Count	28	13	23	64
	%	100.0%	100.0%	100.0%	100.0%

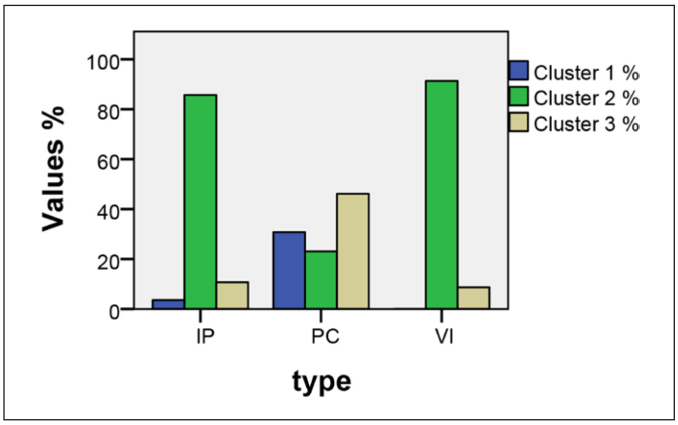

Figure 12: Test 1: 3 clusters, correlation of cluster and type, in %.

interpreted in light of the subtypes (IP, PC, VI) described above, and are thus helpful in validating observed patterns.

A *k*-cluster analysis was run on the final syllables (right boundary) of all tokens in our datasets, with variables corresponding to prosodic correlates (mean F0, excursion size, target slope, target height, target strength, duration, and intensity) as input. A first test specified three clusters, for which the variables were found to be significantly different; the correlates of duration, excursion size, target height and target strength were the prosodic features most representative of membership in each cluster. Two of the three clusters have quite low membership (5 in cluster 1, 48 in cluster 2, and 11 in cluster 3). The cluster labels were then merged with the original dataset so that the distribution of the clusters could be compared with our initial types (IP, PC, VI). A Chi-squared test indicated a significant correlation (χ^2 (4, n = 64) = 25.122, p = 0). The results are shown in Table 7 and graphically in Figure 12, which expresses cluster membership for each subtype as percentages. Clearly, IPs prefer cluster 2, PCs are spread over the three clusters, and VIs also prefer cluster 2.

Because of the low number of tokens in cluster 1 and the correspondence between VIs and IPs in cluster 2, we conducted a second test, this time positing only two groups. The centres for each cluster are shown in Table 8; cluster 1 has larger pitch excursions and more movement in its slope; the target height is lower and has less strength; it is also longer. The results of the ANOVA test for each correlate are shown in Table 9; all are significant, except for mean F0 and intensity.

Table 8: Test 2: Final cluster centres.

	Cluster 1	Cluster 2
mean F0	139.21	155.25
excursion size	6.62	2.57
target slope	–15.80	–11.61
target height	–1.66	4.56
target strength	48.73	61.76
duration	332.34	143.51
intensity	75.18	74.28

Table 9: Statistical tests for the input correlates for the 2 clusters K-means analysis.

Dependent Variable	F	df	Sig.
mean F0	3.027	1, 62	.087
excursion size	24.130	1, 62	.000
target slope	3.137	1, 62	.05
target height	6.342	1, 62	.014
target strength	5.307	1, 62	.025
duration	213.403	1, 62	.000
intensity	.407	1, 62	.526

Table 10: Test 1: 2 clusters, correlation of cluster and type, counts and percentages.

<table>
<tr><th colspan="2" rowspan="2">Cluster</th><th colspan="3">Type</th><th></th></tr>
<tr><th>IP</th><th>PC</th><th>VI</th><th>Total</th></tr>
<tr><td rowspan="2">1</td><td>Count</td><td>3</td><td>10</td><td>2</td><td>15</td></tr>
<tr><td>%</td><td>10.7%</td><td>76.9%</td><td>8.7%</td><td>23.4%</td></tr>
<tr><td rowspan="2">2</td><td>Count</td><td>25</td><td>3</td><td>21</td><td>49</td></tr>
<tr><td>%</td><td>89.3%</td><td>23.1%</td><td>91.3%</td><td>76.6%</td></tr>
<tr><td rowspan="2">Total</td><td>Count</td><td>28</td><td>13</td><td>23</td><td>64</td></tr>
<tr><td>%</td><td>100%</td><td>100%</td><td>100%</td><td>100%</td></tr>
</table>

The Chi-square test shows a significant relation between the clustering membership variable and the types (χ^2 (2, n = 64) = 26.036, p = 0), the results of the counts and percentages are shown in Table 10 and Figure 13.

Although there is variation and not all tokens of one type fall into a given cluster, there are discernible patterns: IPs and VIs group together in cluster 2 and PCs are found predominantly in cluster 1. Given that we are using less-controlled non-lab speech, such variation should not be surprising, the tokens in our datasets vary in their content, being produced by different speakers in different contexts. That patterns emerge in the prosodic correlates is convincing enough, therefore cluster membership is used to test whether unsure cases belong to one or the other category.

Example (10) with accompanying pitch track (Figure 14) is illustrative; it is impossible to decide whether the NP 'the old man' should be interpreted as clause internal or external based on morphological, syntactical and information structural criteria alone.

(10) de_pictures_nlg_005
Na tonikama … [te hola-a]$_{Pred}$ NA HALILI.
ART old.person 3SG.NFUT take-3SG.O ART hook
(CW: Who took what?)
'**The old man**(, **he**) took THE HOOK.'

There is a rather long pause between the fronted constituent and the rest of the clause, which would point to a PC interpretation, but the emphatic marker =*ke* that is so frequently found on constituents in PC position is absent, and there is no co-referential material later on in the clause that would disambiguate this example. As for information structure, it is clear from the context that the constituent is topical. This utterance was made in the context of the picture task (see Section 3.2). Two referents were explicitly introduced, and the speaker was asked who got what from the store. The context and the question invite a parallel contrast construction as an answer, but of course it is not the only available option to the speaker. Both an analysis as one part of a parallel contrast construction (with a contrastive topic in IP position) and an analysis as a normal topic-comment construction (with a shifted/contrastive topic in PC position, and zero encoding of the subject in the clause) are possible and would be appropriate.

Figure 15 shows a comparison of the measurements for example (10) (right) with those of the clear examples (left).

There is a slight fall in pitch from penultimate to final syllable in example (10), similar to that of the IP (blue). The excursion size does not increase substantially from the

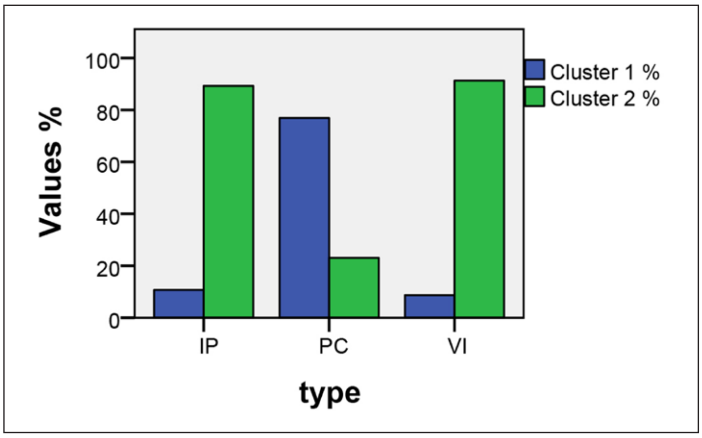

Figure 13: 2 clusters, correlation of cluster and type, in %.

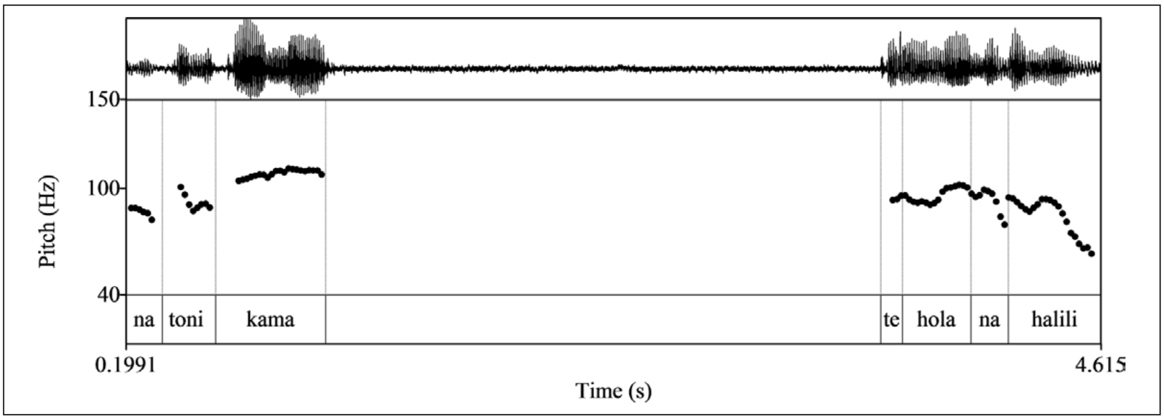

Figure 14: Pitch track of example (10).

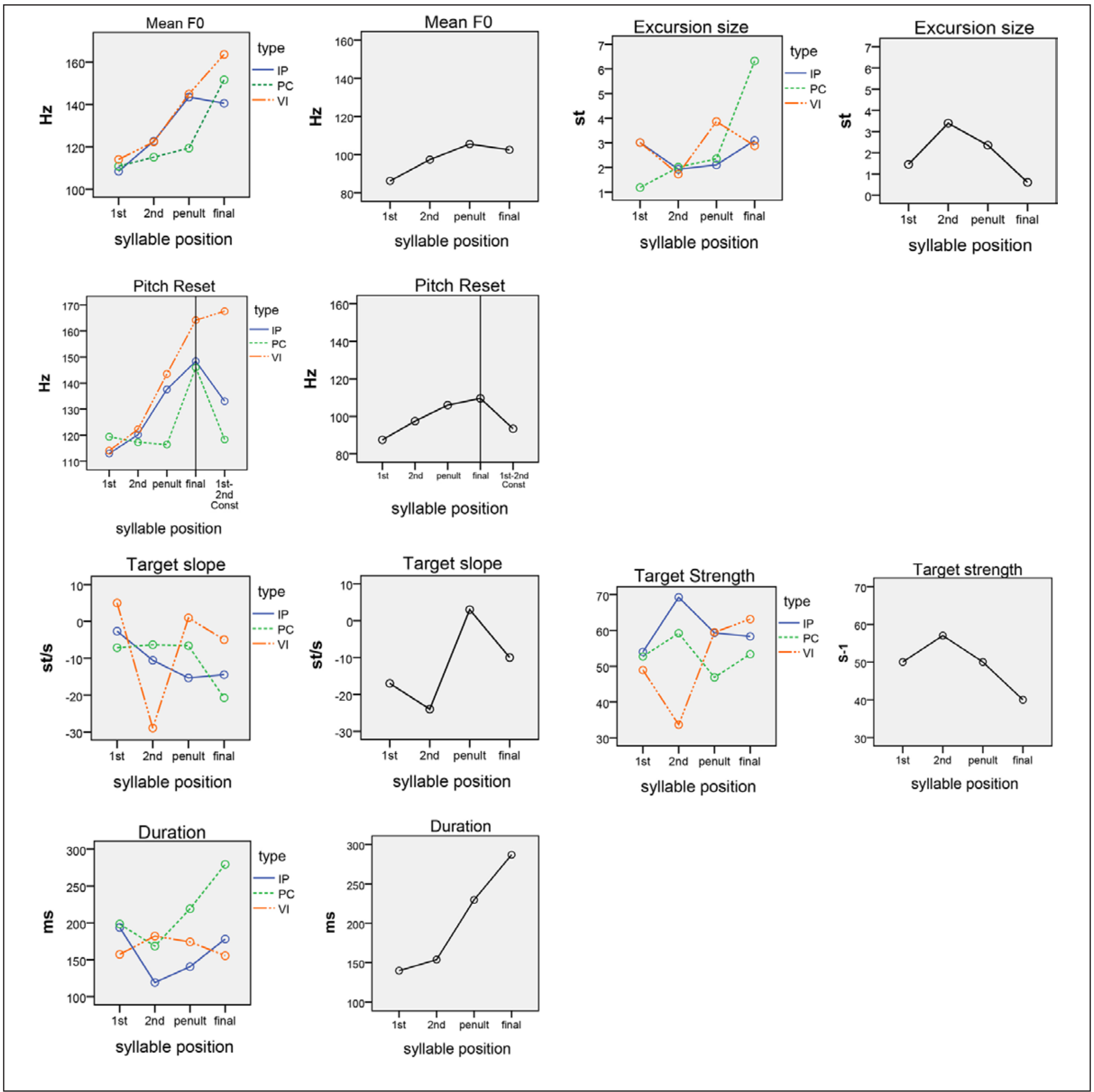

Figure 15: Comparison of the prosodic correlates of example (10) with previous results.

penultimate syllable to the last as for the PC clear cases; instead it decreases, similar to the IP clear cases. There is a pitch reset, from 109.54Hz in the last syllable of the first constituent to 93.39Hz in the first of the second.

The measurements of target slope and strength display patterns that are similar to those of IPs at the right boundary: there is a slight change in the slope from the penultimate to the final syllable; and the measurements are in the same ranges. The strength decreases from the penultimate to the final syllable, but the values are lower than the averages for the IPs.

The durational cue indicates final lengthening, and in this case resembles more the measures of the PC. This forces us to re-evaluate our interpretation which was so far leaning towards an IP, as longer pause durations are important for identifying PCs. Yet it is possible that for this example, the longer pause is an artefact of the elicitation task.

By adding the token to the datasets and running a *k*-means clustering test, it was found that example (10) is a member of cluster 2, mostly associated with IPs. The qualitative and quantitative analyses both support interpreting example (10) as an IP (initial position in the clause), and thus as an example of the parallel contrast construction, mainly based on the correlates associated with pitch, suggesting a [high] pitch target at the right edge, even if there is a long pause.

Example (11) (pitch contour in Figure 16) is also labelled as "unsure". It contains the emphatic particle =*ke* that is frequently found on constituents in PC position, but no pause, and no co-referential material within the remaining clause. The referent of the fronted constituent, the water, was introduced by the speaker earlier in the story; it can be clearly identified as a topic in this utterance. After uttering (11), thus providing the name of the water, he continues to say that the ritual performed there was also called *Vaivari Ao*. Given this context, and similar to example (10), both a PC and IP analysis would make sense.

(11) it_taulaghi_010
Na **beti** **ke** [TARA HOLO-A NI-A]$_{\text{Pred}}$
ART water EMPH 3PL.NFUT call-3SG.O APPL-3SG.O
I VAIVARI AO.
ART Vaivari Ao
(marriage ritual, girls stand in water...)
'**The water** THEY CALL (**it**) VAIVARI AO.'

The measurements are shown in Figure 17. The mean F0 and the slope are closer to those observed in the PC, that is, there is a sharp F0 rise at the right edge, and a falling slope, although values in example (11) are much lower in their range. The pitch reset and duration cue both indicate a phrase boundary. In the *k*-cluster analysis, example (11) is a member of cluster 1, the preferred cluster for PCs.

The measurements and cluster analysis suggest that example (11) is better analysed as a PC (pre-clausal shifted/contrastive topic), as the pitch target at the right edge of the NP is a [rise] rather than a [high] as in example (10).

We set out to test whether prosodic features could help with analysing the syntax of ambiguous constructions in Gela. Firstly, we found that prosodic and syntactic phrases, as far as the constructions under scrutiny are concerned, do align. We also found that the strength of syntactic and prosodic boundaries do correlate, as has been reported for other languages. While it is reputedly not easy to identify boundary correspondences between prosodic and syntactic constituents on the basis of phonetic correlates in non-lab speech (Savy & Voghera 2010), our instrumental analysis has demonstrated that for the clear instances of IPs and PCs, different prosodic patterns can be identified which allow for distinguishing them from each other, as well as from the unmarked VIs. These helped in

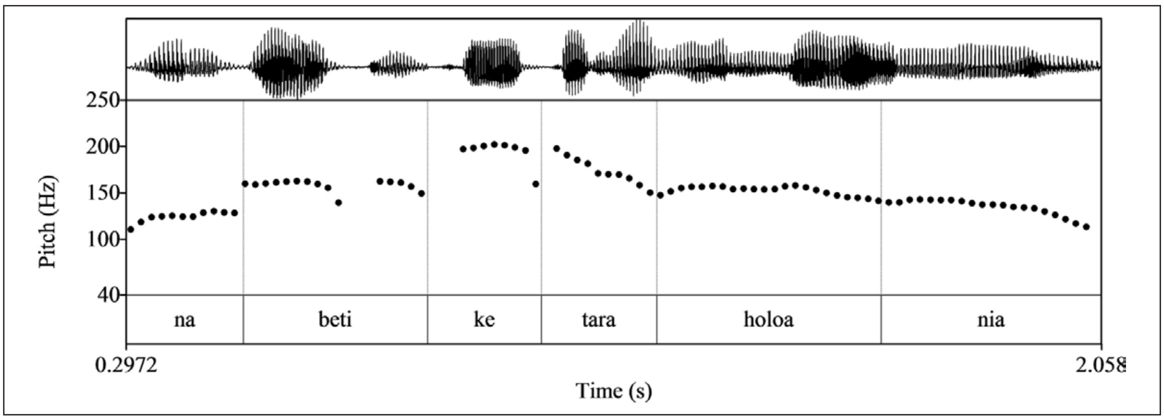

Figure 16: Pitch contour of example (11).

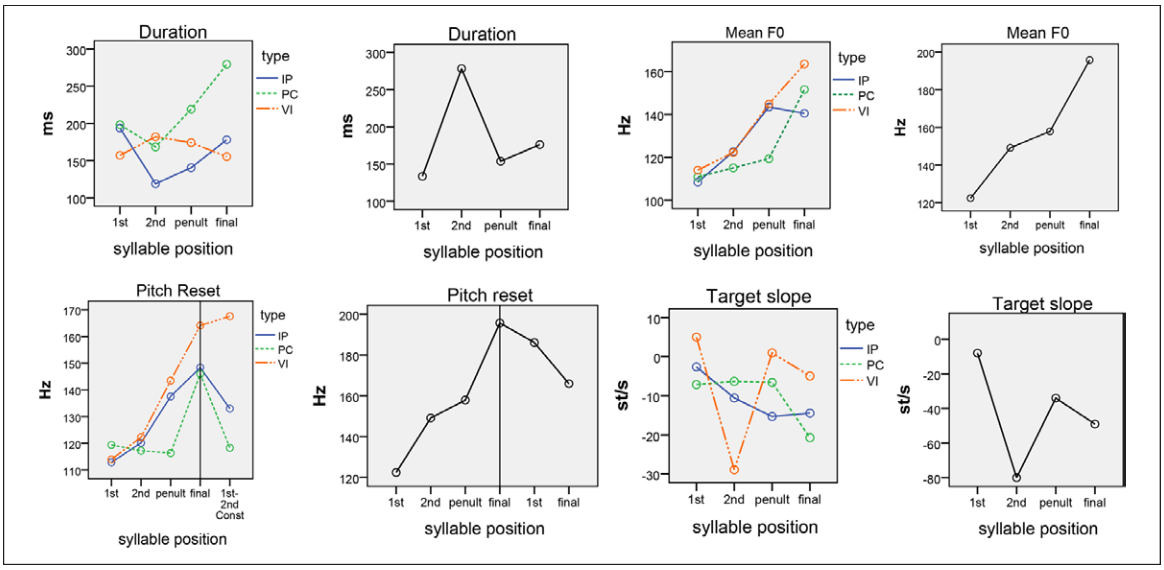

Figure 17: Comparison of the prosodic correlates of example (11) with previous results.

clarifying the status of some of the ambiguous cases as we found that, while not all measurements gave clear results, 12 of the 20 selected tokens exhibited prosodic characteristics that could lead to their being associated more closely with one or the other pattern, leaving a number of cases where the ambiguity is not resolved. This raises an interesting question: are other cues used to resolve this ambiguity, or could it be that this structural ambiguity is not problematic for Gela speakers? That, in context, the syntactic status of fronted NPs can remain unclear? This would pose a challenge for theory. In any case, for the purpose of linguistic description, it would not be possible to claim that the prosodic criteria we have identified can independently distinguish between IP and PC constructions, but the prosodic parameters, in conjunction with information structure and morphosyntax, provide the best set of identifiers.

These findings support the view that prosodic information can be important for spoken language comprehension and especially for syntactic parsing, because prosodic cues can be used to guide the hearer's syntactic analysis. Of course, whether or not hearers actually use all available cues is an open question and requires further research. In addition, our findings suggest that although it is generally acknowledged that there is no one-to-one mapping between prosody and syntax, at least some aspects of syntactic information can be deducible from prosodic contours.

6 Conclusion

This paper introduced new data on prosodic phrasing in Gela, showing that all pre-verbal NPs (both IP and PC) are prosodically separated from the verb (i.e. they constitute a phrase of their own, as can be seen for example in the lengthening on the final syllable), and are in turn distinguished by:

1. Boundary strength: boundary is stronger for IPs, as indicated by length of pauses.
2. Pitch:
   - IP: typically display a [high] target at the right boundary.
   - PC: typically display a [rise] target at the right boundary.

In the past, the foundation for understanding basic grammar had to be based on the investigation of words, phrases, and sentences collected and written down on paper.

Technological advances in audio and video recording, acoustic analysis, and digital corpora and corpus software are now making it possible to investigate language in new ways. It is relatively easy to capture the sound of substantial stretches of spontaneous speech in natural, interactive settings. This opens the way to (re-)analysing better-known languages, as well as including analyses of as yet little known languages. We take the view that an approach that attributes pragmatic meaning to intonation and respects prosody as an independent component of the grammar while seeking to understand its interaction with other components is likely to be more illuminating and ultimately more explanatory.

Abbreviations

1 = first person, 3 = third person, APPL = applicative, ART = article, C = consonant, DU = dual, EMPH = emphasis, IP = clause-initial position, IU = intonation unit, NFUT = non-future tense, NP = noun phrase O = object, PC = pre-clausal position, PL = plural, PROX = proximal, PS = predicate – subject constituent order, PST = past tense, SG = singular, SVC = serial verb construction, V = Vowel, VI = verb-initial clause, VOA = verb – transitive object – transitive subject constituent order, VS = verb – intransitive subject constituent order.

Acknowledgements

The authors would like to thank the Gela and Savosavo speakers for their continuing collaboration in documenting their languages, and for making the recordings available for linguistic research. We thank our team of research assistants for assistance with annotating and preparing the data for acoustic analysis.

We would also like to thank the audience of the ETI3 in Montreal and the Linguistics Colloquium at the University of Bielefeld for helpful comments. We thank the three anonymous reviewers whose comments and suggestions helped improve and clarify this paper.

This research paper is part of the Project "Discourse and Prosody across Language Family Boundaries", funded by the Volkswagen Foundation (PA 86101). Part of Wegener's Gela data was collected with funding from the Max Planck Research Group for Comparative Population Linguistics at the MPI for Evolutionary Anthropology, Leipzig.

Competing interests

The authors declare that they have no competing interests.

References

Bauer, Winifred. 1991. Māori *ko* again. *Te Reo* 26. 3–14.

Boersma, Paul & David Weenink. 2013. Praat: Doing phonetics by computer. Retrieved from http://www.praat.org/on2013/11/30.

Bögel, Tina, Miriam Butt, Ronald M. Kaplan, Tracy Holloway King & John T. Maxwell III. 2009. Prosodic phonology in LFG: A new proposal. *Proceedings of the LFG2009 Conference*. csli-publications.stanford.edu/LFG/14/lfg09.html.

Büring, Daniel. 1997. *The meaning of topic and focus*. London: Routledge. DOI: https://doi.org/10.4324/9780203448014

Calhoun, Sasha. 2015. The interaction of prosody and syntax in Samoan focus marking. *Lingua*165 B. 205–229. DOI: https://doi.org/10.1016/j.lingua.2014.11.007

Chafe, Wallace. 1987. Cognitive constraints and information flow. In Russell Tomlin (ed.), *Coherence and grounding in discourse: Outcome of a symposium*, 21–51. Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/tsl.11.03cha

Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Clemens, Lauren Eby. 2014. *Prosodic noun incorporation and verb-initial syntax*. Harvard, MA: Harvard University dissertation. http://nrs.harvard.edu/urn-3:HUL.InstRepos:13070029.

Clemens, Lauren Eby & Maria Polinsky. In press. Verb-initial word orders (primarily in Austronesian and Mayan languages). In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*. 2nd Edition; Oxford: Blackwell Publishers.

Cook, Philippa & Felix Bildhauer. 2011. Annotating Information Structure: The case of topic. In Stefanie Dipper & Heike Zinsmeister (eds.), *Beyond semantics. Corpus-based investigations of pragmatic and discourse phenomena*, 45–56. Bochum: Universität Bochum, Bochumer Linguistische Arbeitsberichte.

Cooper, William E. & Jeanne Paccia-Cooper. 1980. *Syntax and speech*. Cambridge, MA: Harvard University Press. DOI: https://doi.org/10.4159/harvard.9780674283947

Crowley, Terry. 2002. Gela. In John Lynch, Malcolm Ross & Terry Crowley (eds.), *The Oceanic languages*, 525–537. Richmond: Curzon Press.

D'Imperio, Mariapaola, Gorka Elordieta, Sonia Frota, Pilar Prieto & Marina Vigário. 2005. Intonational Phrasing in Romance: The role of prosodic and syntactic structure. In Sonia Frota, Marina Vigário & Maria J. Freitas (eds.), *Prosodies,* 59–97. Berlin: Mouton de Gruyter.

Elfner, Emily. 2012. *Syntax-prosody interaction in Irish*. Amherst, MA: University of Massachusetts dissertation. http://scholarworks.umass.edu/open_access_dissertations/505.

Elordieta, Gorka, Sonia Frota, Pilar Prieto & Marina Vigário. 2003. Effects of constituent weight and syntactic branching of phrasing decisions in Romance. *Proceedings of the XIVth International Congress of Phonetic Sciences*. 487–490.

Fanselow, Gisbert. 2007. The restricted access of information structure to syntax: A minority report. In Caroline Féry, Gisbert Fanselow & Manfred Krifka (eds.), *The notions of information structure* (Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure (ISIS) 6), 205–220. Potsdam: Universitätsverlag Potsdam.

Fodor, Janet Dean. 2002. Prosodic disambiguation in silent reading. In *Proceedings of the North East Linguistic Society* 32. 113–132. Amherst, MA: GSLA, University of Massachusetts.

Foley, William. 1986. *The Papuan languages of New Guinea.* Cambridge: Cambridge University Press.

Foley, William A. & Robert D. Van Valin, Jr. 1985. Information packaging in the clause. In Timothy Shopen (ed.), *Language Typology and Syntactic Description, Vol. I: Clause Structure*, 282–364. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511815515

Gee, James Paul & François Grosjean. 1983. Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology* 15(4). 411–458. DOI: https://doi.org/10.1016/0010-0285(83)90014-2

Givón. Talmy. 1983. Topic continuity in discourse: An introduction. In Talmy Givón (ed.), *Topic continuity in discourse*. Amsterdam/Philadelphia: John Benjamins Publishing Co. DOI: https://doi.org/10.1075/tsl.3

Goldberg, Adele. 1995. *Constructions: A Construction Grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, Adele. 2006. *Constructions at work*. Oxford: Oxford University Press.

Goldman-Eilser, Frieda. 1968. *Psycholinguistics: Experiments in spontaneous speech*. New York, NY: Academic Press. DOI: https://doi.org/10.1016/0024-3841(70)90028-8
Herring, Susan C. 1990. Information structure as a consequence of word order type. *Proceedings of the 16th Annual Meeting of the Berkeley Linguistics Society*. 163–174. DOI: https://doi.org/10.3765/bls.v16i1.3363
Hill, Deborah. 1992. *Longgu grammar*. Canberra: Australian National University dissertation.
Hill, Deborah. 2002. Longgu. In John Lynch, Malcolm Ross & Terry Crowley (eds.), *The Oceanic languages*, 538–561. Richmond: Curzon Press.
Jepson, Kathleen. 2014. Intonational marking of focus in Torau. *Selected Papers from the 44th Conference of the Australian Linguistics Society*. Melbourne: University of Melbourne.
Jung, Sun-Ah. 2003. The effect of phrase length and speech rate on prosodic phrasing. In Maria-Josep Solé, Daniel Recasens & Joaquin Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, Vol. 1. 483–486.
Klein, Ewan. 2000. Prosodic constituency in HPSG. In Ronnie Cann, Claire Grover & Phillip H. Miller (eds.), *Grammatical interfaces in HPSG*, 169–200. CSLI Publications.
Krifka, Manfred. 2007. Basic notions of Information Structure. In Caroline Féry, Gisbert Fanselow & Manfred Krifka (eds.), *The notions of Information Structure* (Working Papers of the SFB 632. Interdisciplinary Studies on Information Structure (ISIS) 6), 13–56. Potsdam: Universitätsverlag Potsdam.
Ladd, D. Robert. 1996. *Intonational Phonology* (Cambridge Studies in Linguistics 79). Cambridge: Cambridge University Press.
Lambrecht, Knud. 1994. *Information structure and sentence form. Topic, focus and the mental representations of discourse referents*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511620607
Lehiste, Ilse. 1983. The many linguistic functions of duration. *The Journal of the Acoustical Society of America* 73. S65. DOI: https://doi.org/10.1075/cilt.32.11leh
Lichtenberk, Frantisek. 2008. *A grammar of Toqabaqita* (Mouton Grammar Library 42). Berlin: Mouton de Gruyter.
Longacre, Robert. 1995. Left shifts in strongly VSO languages. In Pamela Downing & Michael Noonan (eds.), *Word order and discours.* Amsterdam: Benjamins. DOI: https://doi.org/10.1075/tsl.30.12lon
Lynch, John, Malcolm Ross & Terry Crowley. 2002. *The Oceanic languages.* Richmond: Curzon Press.
Maslova, Elena & Giuliano Bernini. 2006. Sentence topics in the languages of Europe and beyond. In Giuliano Bernini & Marcia L. Schwartz (eds.), *Pragmatic organization of discourse in the languages of Europe. Empirical approaches to language typology* 20–28, 67–120. Berlin/New York: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110892222.67
Miller, Ingrid. 1974. Gēla syntax. London: University of London, School of Oriental and African Studies dissertation. http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.527018.
Nolan, Francis. 2003. Intonational equivalence: An experimental evaluation of pitch scales. In *Proceedings of 15th ICPhS.* 771–774. Barcelona, Spain.
Palmer, Bill. 2009. Clause order and information structure in Cheke Holo. *Oceanic Linguistics* 48(1). 213–249. DOI: https://doi.org/10.1353/ol.0.0038
Payne, Doris. 1999. What counts as explanation? A Functionalist approach to word order. In Michael Darnell, Edith Moravcsik, Frederick J. Newmeyer, Michael Noonan & Kathleen Wheatley (eds.), *Functionalism and Formalism in Linguistics*, Vol. 1, 135–163. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/slcs.41.08pay

Pederson, Eric, Eve Danziger, David Wilkins, Stephen Levinson, Sotaro Kita & Gunter Senft. 1998. Semantic typology and spatial conceptualization. *Language* 74(3). 557–589. DOI: https://doi.org/10.2307/417793

Prom-on, Santitham, Fang Liu & Yi Xu. 2011. Functional modeling of tone, focus, and sentence type in Mandarin Chinese. In *Proceedings of the 17th ICPhS*. 1638–1641.

Prom-on, Santitham, Yi Xu & Bundit Thipakorn. 2009. Modeling tone and intonation in Mandarin and English as a process of target approximation. *Journal of the Acoustical Society of America* 125. 405–424. DOI: https://doi.org/10.1121/1.3037222

Repp, Sophie. 2016. Contrast: Dissecting an elusive information-structural notion and its role in grammar. In Caroline Féry & Shinichiro Ishihara (eds.), *OUP handbook of Information Structure,* 270–289. Oxford: Oxford University Press.

Rooth, Mats. 1985. *Association with focus*. Amherst, MA: University of Massachusetts dissertation. http://scholarworks.umass.edu/dissertations/AAI850959.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1. 75–116. DOI: https://doi.org/10.1007/BF02342617

Savy, Renata & Miriam Voghera. 2010. A corpus-based study on syntactic and phonetic prosodic phrasing boundaries in spontaneous Italian speech. In *Proceedings of ISCA Speech Prosody 2010*. 1–4. Chicago.

Slioussar, Natalia. 2007. *Grammar and information structure. A study with reference to Russian*. Utrecht: LOT Publications dissertation.

Solomon Islands National Statistical Office. 2009. *Report on 2009 Population and Housing Census: Basic Tables and Census Description,* Vol. 1: Honiara: Solomon Islands National Statistical Office.

Szendrői, Kriszta. 2003. A stress-based approach to the syntax of Hungarian focus. *The Linguistic Review* 2. 37–78. DOI: https://doi.org/10.1515/tlir.2003.002

Turk, Alice E. & Stephanie Shattuck-Hufnagel. 2007. Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics* 35. 445–72. DOI: https://doi.org/10.1016/j.wocn.2006.12.001

Unger, Paul. 2008. *Aspects of Lengo grammar.* Langley: Trinity Western University, Canada Institute of Linguistics MA Thesis.

Van Valin, Robert D. Jr. 1993. A synopsis of role and reference grammar. In Robert D. van Valin Jr. (ed.), *Advances in role and reference grammar*, 1–164. Amsterdam: John Benjamins. DOI: https://doi.org/10.1017/CBO9781139166799

Van Valin, Robert D. Jr. & Randy LaPolla. 1997. *Syntax: Structure, meaning and function.* Cambridge: Cambridge University Press.

Wagner, Michael. 2005. *Prosody and recursion.* Cambridge, MA: Massachusetts Institute of Technology dissertation. http://hdl.handle.net/1721.1/33713.

Wegener, Claudia. In prep. Basic constituent order and information-structural functions of constituent order variation in Savosavo (Papuan) and Gela (Oceanic) corpus data.

Wightman, Colin W., Stephanie Shattuck-Hufnagel, Mari Ostendorf & Patti J. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustic Society of America,* Vol. 91(3), 1707–17. DOI: https://doi.org/10.1121/1.402450

Xu, Yi. 2005. Speech melody as articulatorily implemented communicative functions. *Speech Communication* 46. 220–251. DOI: https://doi.org/10.1016/j.specom.2005.02.014

Xu, Yi. 2009. Timing and coordination in tone and intonation – An articulatory-functional perspective. *Lingua* 119. 906–927. DOI: https://doi.org/10.1016/j.lingua.2007.09.015

Xu, Yi. 2013. ProsodyPro: A tool for largescale systematic prosody analysis. In *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013*).

Xu, Yi, Albert Lee, Santitham Prom-on & Fang Liu. 2015. Explaining the PENTA model: A reply to Arvaniti and Ladd. *Phonology* 32(3). 505–535. DOI: https://doi.org/10.1017/S0952675715000299

Xu, Yi & Santitham Prom-on. 2014. Toward invariant functional representations of variable surface fundamental frequency contours: Synthesizing speech melody via model-based stochastic learning. *Speech Communication* 57. 181–208. DOI: https://doi.org/10.1016/j.specom.2013.09.013

Zec, Draga & Sharon Inkelas. 1990. Prosodically constrained syntax. In Sharon Inkelas & Draga Zec (eds.), *The phonology-syntax connection*, 365–378. Chicago: University of Chicago Press.

How to cite this article: Simard, Candide and Claudia Wegener. 2017. Fronted NPs in a verb-initial language – clause-internal or external? Prosodic cues to the rescue!. *Glossa: a journal of general linguistics* 2(1): 51. 1–32, DOI: https://doi.org/10.5334/gjgl.211

Published: 16 May 2017

]u[*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS