The syntax-prosody interface in LFG: Revisiting Korean question focus

Stephen M. Jones
University of Groningen

Yoolim Kim
Wellesley College

Cong Zhang
Newcastle University

Proceedings of the LFG'24 Conference

Miriam Butt, Jamie Y. Findlay and Ida Toivonen (Editors)

2024

PubliKon

lfg-proceedings.org

Abstract

This paper sets out two challenges to LFG analyses of the syntax-prosody interface—one general, one specific—arising from experimental evidence that shows considerable hearer variation in the interpretation of specific prosodic patterns in Korean that are canonically associated with the scope of question focus. We assume that this arises from a lexical preference for the question word reading of content pro-forms that are ambiguous between question words and indefinite pronouns, which has an effect on how pitch contours are perceived. A revised formal treatment of the phenomenon, building on Jones (2016) is presented, which robustly handles the variation in hearer perception and links to ongoing work to model hearers' decision-making process.

1 Introduction

Prosody is widely held to play a crucial role in disambiguating different readings of identical strings in Korean, including differentiating between statements and questions, and distinguishing content from polar questions.[†] However, new evidence (Jones et al. under review) suggests that the canonical account of Korean prosodic disambiguation does not always hold. An experiment that started by trying to more precisely specify the amount of F0 variation required for a hearer to register focus, turned out to open up a wider question regarding the assumptions on ambiguity. In this paper we propose an analysis that revises the definition of the prosodic characteristics of question focus given in Jones (2016), and which is amenable to variation in hearer comprehension.

In the following sections, we give a brief overview of the phenomenon, of various accounts of the prosodic characteristics used to disambiguate, and of LFG accounts that incorporate prosody into analyses. We then present experimental evidence before detailing the revised account.

2 Background

Written sentences in the Korean polite speech style, marked by the sentence-final particle *-yo*, are ambiguous between declarative, imperative, interrogative and propositive moods. Canonically, this ambiguity disappears in the spoken language (see e.g. Jun & Oh 1996). Specific utterance-final tunes distinguish between declarative and interrogative mood: Jun (2005) describes declarative mood as being marked by a HL% tune[1], whereas interrogative mood is marked with an LH% tune. Within interrogative mood, there is consensus that content and polar questions are distinguished prosodically, but the exact nature of these differences is still the subject of investigation and debate.

When combined with content pro-forms (CPFs)—words that are ambiguous between content question words (wh-words) and indefinite pro-forms—the resulting sentence has statement, content question, and polar question readings (1).

[†]We thank Jacolien van Rij for her assistance and advice in building the statistical models. We also thank two anonymous reviewers and the editors, especially Jamie Findlay, for their comments and suggestions, which have substantially improved the analysis.

[1]In the description of tones, H and L stand for high and low tones respectively. % indicates the boundary of an utterance, so HL% is an utterance-final HL tone sequence, or tune.

(1) *hakchangsicel ttay nwukwu-lul mollay sarangh-aysse-yo*
school.days during who/someone-OBJ secretly love-PST-POL

a. "You secretly loved someone when you were at school."[2] (statement)
b. "Who did you secretly love when you were at school?" (content question)
c. "Did you secretly love someone when you were at school?" (polar question)

2.1 Korean prosody

Jun (2005) gives an account of Korean prosody in which prosodic phrases are marked by characteristic tone patterns at their left and right edges, with phrase-final lengthening. Jun's model includes syllables, prosodic words, accentual phrases (AccP) and intonational phrases (IntP). For this analysis, the AccP and IntP are most important: Jun claims that prosodic words are unmarked, with no evidence for lexical stress. Within prosodic structure, AccPs are nested inside IntPs in line with Selkirk's (1984) Strict Layer Hypothesis. However, the boundaries of prosodic constituents are not constrained by syntactic constituent boundaries: a syntactic phrase with many syllables may be split between AccPs, and these AccPs may also cover parts of adjacent syntactic phrases.

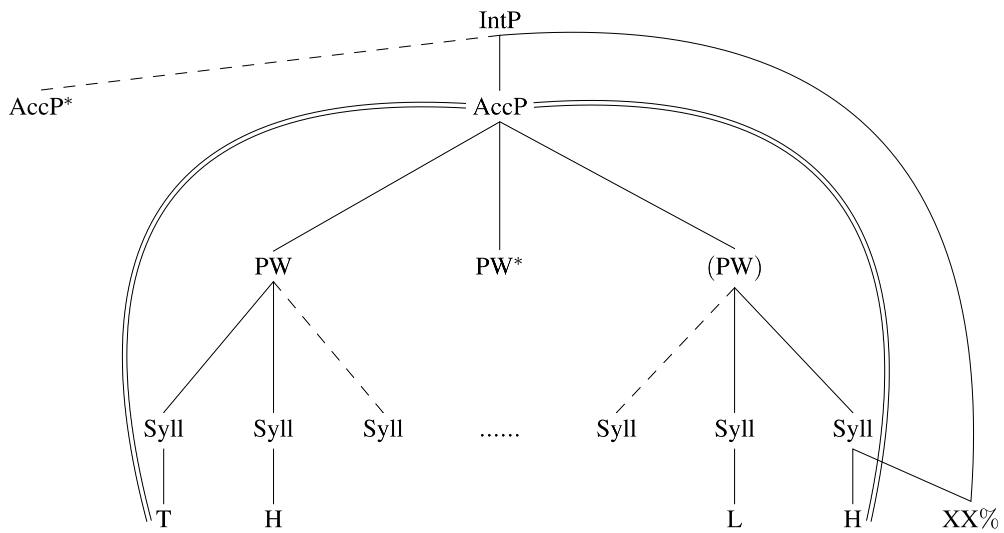

Figure 1: General intonation structure of Korean, including prosodic hierarchy, proposed by Jun (2005: 205). T represents either an H or L tone, phonologically conditioned by the onset of the first syllable. An IntP-final boundary tune, represented here as XX%, replaces the final H of the last AccP in an IntP

2.2 Accounting for the phenomenon

Accounts of the phenomenon rest on two assumptions: a difference in scope of focus between the two question types, and a characteristic prosodic pattern that is associated

[2] Glossing abbreviations follow the Leipzig Glossing Rules apart from POL, which indicates the informal polite speech style. In all three translations, the subject of the utterance is unspecified and pragmatically determined; we have given 'you' as an example.

with the presence of focus. Following Dalrymple & Nikolaeva (2011) we assume that polar questions have broad focus: the whole of the proposition expressed by the question is at issue. Content questions show narrow focus: the CPF constrains the set of felicitous answers and so bears the question focus.

The various accounts of the disambiguation rest on a combination of prosodic characteristics including AccP boundaries (marked by characteristic tone patterns but not necessarily breaks), F0 values associated with focus, post-focus pitch compression, and the presence or absence of L tones within AccPs. The earliest account is that of Jun & Oh (1996), who propose that differences in the placement of AccP boundaries in relation to the CPF and the verb are the primary determiner, albeit with some observed differences in the nature of the final tone/tune (varying between H% or LH%).

Jones (2016) carried out a speech production experiment and observed raised F0 pitch spreading left from the right edge of the constituent holding question focus: within the CPF for content questions and at the verb for polar questions. Based on this, he proposed a prosodic feature EXPANDED PITCH RANGE.

In contrast to these approaches, Yun (2019) identified dephrasing after the CPF as the primary disambiguator. In experimental manipulations this was found to be more influential than the size of the F0 pitch peak at the CPF. Further experiments (Yun & Lee 2022) resulted in a more fine-grained view, identifying a larger number of contributory factors. A raised F0 pitch peak at the CPF was sufficient, but not necessary, for a question to be interpreted as a content question. For content questions, changing the final tone from LH% to H% increased the likelihood of the question being perceived as polar. And for polar questions, changing the final tone from H% to LH% and removing the L tone after the CPF increased the chance of being perceived as a content question.

2.3 LFG and prosody

LFG is capable of including prosodic structure within its modular framework. There are two current approaches: in this paper we are using the approach proposed by Dalrymple & Mycock (2011) and elaborated by Mycock & Lowe (2013). Details of the other approach (Bögel 2015, 2022) and a more detailed discussion of the history of LFG accounts of the syntax-prosody interface can be found in Bögel (2023).

In the approach first set out by Dalrymple & Mycock (2011) and subsequently elaborated by Mycock & Lowe (2013) and Dalrymple et al. (2019: 406 ff.), prosody and syntax are connected by a mutually constraining relationship between p-string elements (which have a functional relationship with the terminal nodes of p-structure) and s-string elements (which have a functional relationship with the terminal nodes of c-structure). All p- and s-string elements are assumed to have features L and R. Where the associated terminal nodes in p- or c-structure sit at the left or right edge of a constituent at any level of their containing structures, that information is captured in the sets that are the values of features L and R respectively.

Edge sets can also hold information about the prosodic or syntactic expression of information-structural features such as focus. Where a syntactic constituent corresponds to the scope of focus, the node representing that constituent holds a feature representing the exponence of syntactic focus. Language-specific cascade rules determine how the exponent feature is passed down the c-structure tree into the terminal nodes, such that it appears in an edge set of the left or right edge of the constituent. The mutual constraints between prosody and syntax are governed by the principle of Interface Harmony. This states that an utterance is well-formed if and only if the prosodic and syntactic exponents of an information structural attribute are contained in the the corresponding p-structure and s-string edge sets respectively.

In this model, prosodic exponence of focus is captured by rules that render acoustic information as abstract prosodic features. These prosodic features are linked with features representing the prosodic exponence of information-structure attributes such as focus. Again, language-specific rules determine which of the p-structure edge sets a prosodic exponent sits in.

2.4 LFG and Korean prosody

To date, the only attempt to provide an LFG account of the syntax-prosody interface in Korean is Jones (2016). His account followed Dalrymple & Mycock (2011) and Mycock & Lowe (2013) and assumed Jun's (2005) description of Korean prosodic structure. Based on a speech production experiment, he proposed the prosodic feature EXPANDED PITCH RANGE which was minimally present at the right edge of an AccP holding question focus, and which could spread left through the AccP. Variations in the point of onset of EXPANDED PITCH RANGE were seen as due to individual speaker variation, and because the critical edge set was the right edge of the AccP, the variation did not have a material effect on the account.

Jones's (2016) account is unsatisfactory because the nature of the proposed prosodic feature EXPANDED PITCH RANGE is not precisely specified. He presents evidence that shows regular pitch differences within the AccP hosting question focus, but does not address the the question of precisely what size of pitch range is needed for EXPANDED PITCH RANGE to be taken as present, compared with F0 variation in AccPs that do not exhibit the prosodic feature.

3 Experimental evidence

In an attempt to sharpen the definition of EXPANDED PITCH RANGE, the authors carried out a large scale ($n = 124$) online speech perception study. The experiment followed the gating paradigm, using stimuli where size of the natural F0 range had been artificially reduced in the regions of interest.

3.1 Method

The experimental task was to identify whether experimental stimuli were statements, content questions or polar questions, or whether it was not possible to tell. Twenty-one natural stimulus sets were generated from three-way ambiguous strings (statement/content question/polar question). These sets were then recorded in each of the variants with natural prosody by a native speaker of Seoul Korean. The natural stimuli for content and polar questions were then manipulated in Praat (Boersma 2001) using scripts derived from models provided by Lennes (2017) to vary the F0 contour in the region of interest, where Jones (2016) predicted that the prosodic feature EXPANDED PITCH RANGE would be found.

For content questions, the region of interest was the AccP containing the CPF: the difference between minimum and maximum F0 in this AccP was measured, and compared with the difference between the minimum and maximum F0 in the corresponding AccP in the statement. Natural pitch range expansion was then calculated as the difference between these two values. The F0 contour of the question stimulus was smoothed, and the pitch peak manipulated to produce variants with the F0 peak in the region of interest at 25%, 50% and 75% of the natural pitch range expansion.

For polar questions, the region of interest was the final AccP which contains the verb including the utterance-final LH% tune. Here, two pitch ranges were measured: the difference between maximum and minimum F0 during the verb stem, and the difference between maximum and minimum F0 during the the utterance-final tune. The F0 values for the content question were taken as the baseline (0% natural focus prosody), with the difference between the polar question and the content question taken as the natural pitch range expansion. Again, the pitch contours of the polar question stimuli were smoothed, and the pitch maximum was manipulated to create variants with F0 peaks at 25%, 50% and 75% of the natural pitch range expansion. Thus the three stimuli in a natural stimulus set resulted in a manipulated stimulus set of 11 stimuli (5 × content question, 5 × polar question, 1 × statement) for each of the 21 strings. The manipulated stimuli were then tested for acceptability with native speakers, with the mean acceptability score being 4.54 on a scale of 1 (low)–5 (high).

Once the manipulated stimulus sets had been created, all stimuli were segmented into incrementally-longer fragments following the schema in Figure 2. Thus for each stimulus within a stimulus set, there was a test item set of five audio files. The boundaries between segments mostly coincided with word boundaries apart from the boundary between segments 4 and 5, which was the boundary between the verb and the sentence-final particle *-yo*. This was done because the final particle carries the utterance-final tune (HL% for statements and LH% for questions), and we were interested to know whether participants distinguished questions from statements before hearing the categorical prosodic information.

	Segments				
	1	2	3	4	5
a.	*hakchangsicel ttay*				
b.	*hakchangsicel ttay*	*nwukwulul*			
c.	*hakchangsicel ttay*	*nwukwulul*	*mollay*		
d.	*hakchangsicel ttay*	*nwukwulul*	*mollay*	*saranghaysse-*	
e.	*hakchangsicel ttay*	*nwukwulul*	*mollay*	*saranghaysse-*	*yo*
	schooldays during	who/someone	secretly	loved	POL

Figure 2: Rows a.–e. show the iterated, incrementally lengthened presentation of a stimulus. After each presentation, participants were asked to identify the utterance—if they could—as a statement, a content question, or a polar question. The region of interest, where canonical prosodic expression of the scope of focus is found, is in segment 2 for content questions and in segments 4–5 for polar questions

The experiment was hosted on a JATOS server (Lange et al. 2015) and written using jsPsych (de Leeuw et al. 2023). Participants used their own computer equipment for the experiment. Participants were native speakers of Korean recruited using the participant recruitment service Prolific,[3] and were remunerated for their time at the rate of GBP 12 per hour. Test item sets were presented iteratively, with incrementally long fragments. After each fragment was presented, participants were asked to make a judgement on the type of stimulus. A Latin Square design was used so that each participant was exposed to only one item from each manipulated stimulus set, with the 11 variants counterbalanced across participants. Full details of the preparation of the stimuli and the experimental procedure are given in Jones et al. (under review).

Following Jones (2016), we predicted that for stimuli with 100% natural pitch range expansion, content questions would be disambiguated at segment 2 when expanded pitch range was heard at the CPF, and that polar questions would be disambiguated at segment 4, when expanded pitch range was heard at the verb. We also predicted a gradient effect of the reduction in natural pitch expansion, with accuracy decreasing and disambiguation taking place later in the utterance as the proportion of natural pitch expansion decreased.

3.2 Results

The results were analysed using R 4.3.2 (R Core Team 2023) and the packages *lme4* (Bates et al. 2015), *mgcv* (Wood 2017), and *itsadug* (van Rij et al. 2022). Generalised additive models (Hastie & Tibshirani 1990) were constructed separately for content and polar question data, with fixed effects of segment and of the proportion of natural pitch range expansion present, and random effects of participant, item and order of presentation. For each type of question, pairs of models were constructed with and without an interaction between segment and the proportion of natural prosody present. For content questions, the model without an interaction was preferred, whereas for polar questions the model with an interaction was preferred. Further details of the statistical analysis are available in Jones et al. (under review) and the study results suggest that other factors are in play.

Figure 3 shows that statements were successfully disambiguated from questions, and that this took place at the final particle *-yo*, which was marked by the utterance-final tune. This provides further evidence in support of Jun's (2005) widely-accepted description of the HL% and LH% utterance-final tunes as being the defining prosodic characteristic of statements and questions respectively. The relationship between correct disambiguation and distance through the utterance is non-linear: there is a small though not-significant reduction in accuracy at the second segment, which is the CPF.

Summary plots of the preferred statistical models for content and polar questions are shown at Figures 4 and 5 respectively. The findings are not in line with the predictions and so do not support the hypothesis derived from Jones (2016) about the nature of EXPANDED PITCH RANGE.

Figure 4 shows that content question stimuli were ultimately disambiguated successfully. However, this was only at the final segment. There was no successful disambiguation at the CPF, and although the level of successful disambiguation rose as

[3] https://prolific.com

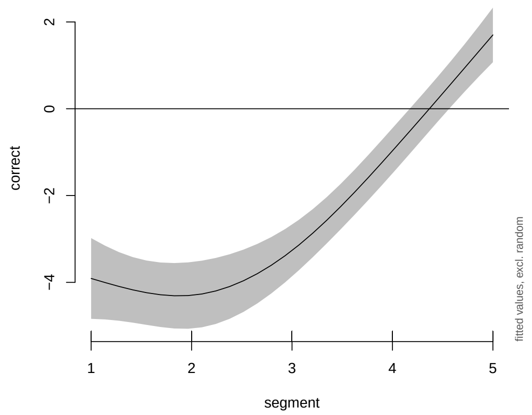

Figure 3: Disambiguation of statement stimuli. X-axis = region of stimulus: 1 = opening constituent; 2 = CPF; 3 = adverbial; 4 = verb; 5 = final particle *-yo*. Y-axis = log odds of successful disambiguation: 0 = chance

participants heard more of the stimuli, it was only at segment 4 that these reached chance level. Throughout the presentation, there was no significant difference in the levels of disambiguation, regardless of the proportion of natural pitch range expansion that was present. Even with 0% of natural focus prosody, with the CPF having the same pitch range as that of a natural polar question, the stimuli were interpreted as content questions.

Figure 5 shows the corresponding results for polar question stimuli. Here, a gradient effect of prosody was seen, with an interaction between the proportion of natural pitch range expansion and segment. For stimuli with full natural pitch range expansion, the chance of successful disambiguation increased more rapidly than for stimuli with 0% of the natural pitch range expansion. However, even where 100% of the natural pitch range expansion was present, correct disambiguation only occurred at chance levels at the final particle *-yo*: thus the speaker's intent to communicate a polar question was inconsistently perceived by participants. For the variants with a lower proportion of natural pitch range expansion, stimuli were always significantly likely to be incorrectly disambiguated.

Figure 6 shows the final response given by participants to polar question stimuli. For levels of natural pitch range expansion below 75%, participants significantly misidentified the stimuli as content questions. For 75% and 100% of natural pitch range expansion, there was no significant difference between the correct identification and misidentification as a content question. In other words, even without the pitch range expansion at the CPF, which is canonically associated with content questions, and in the absence of natural pitch range expansion at the verb, which is canonically associated with polar questions, participants showed a preference to interpret a polar question as a content question. Where pitch range expansion was present at the verb, this was not enough for participants to significantly prefer the canonical interpretation as a polar question.

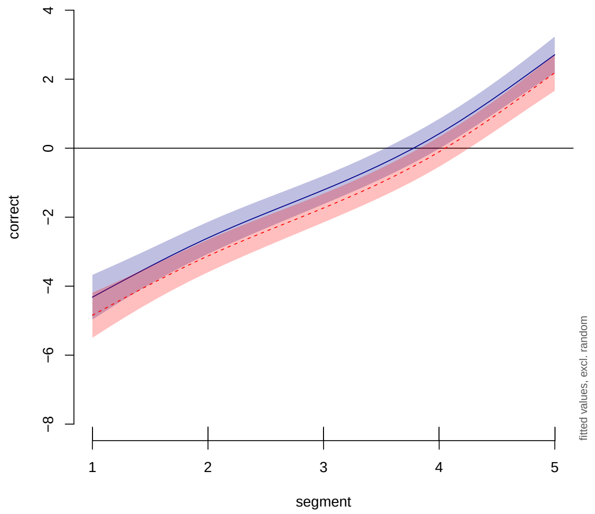

Figure 4: Disambiguation of content question stimuli. X-axis = region of stimulus: 1 = opening constituent; 2 = CPF; 3 = adverbial; 4 = verb; 5 = final particle *-yo*. Y-axis = log odds of successful disambiguation: 0 = chance. Solid blue = 100% and dashed red = 0% of natural pitch range expansion respectively

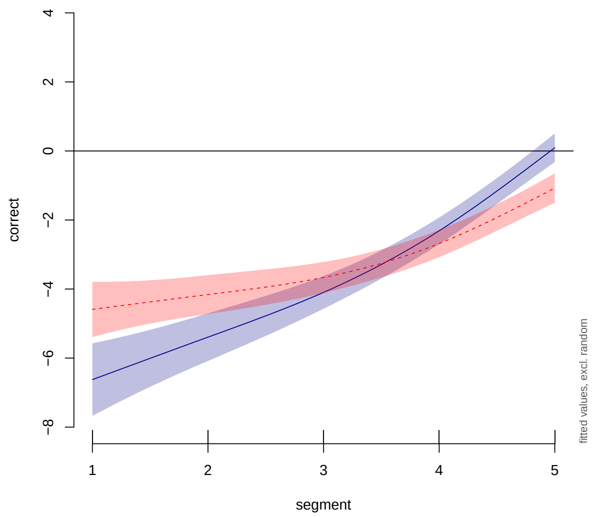

Figure 5: Disambiguation of polar question stimuli. X-axis = region of stimulus: 1 = opening constituent; 2 = CPF; 3 = adverbial; 4 = verb; 5 = final particle *-yo*. Y-axis = log odds of successful disambiguation: 0 = chance. Solid blue = 100% and dashed red = 0% of natural pitch range expansion respectively

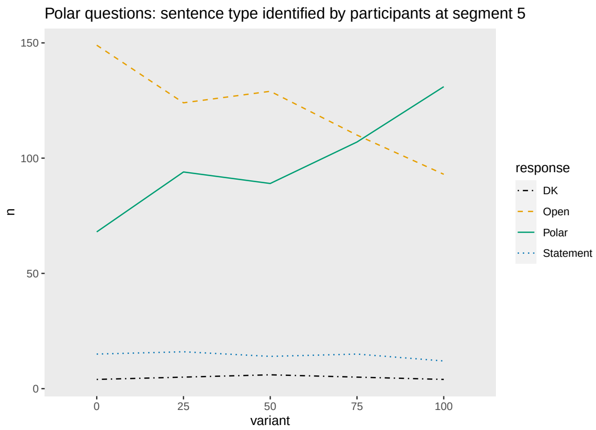

Figure 6: How participants identified polar questions once the full utterance had been heard. X-axis = percentage of natural pitch range expansion. Y-axis = number of responses. Green/solid = correctly identified as a polar question. Yellow/dashed = incorrectly identified as a content question. Blue/dotted = incorrectly identified as a statement. Black/dot-dash = don't know

In a small proportion of trials, less than 10%, participants identified the stimuli as statements: we assume that they did not hear the final LH% tune. For space reasons, detailed analysis of this is excluded from this paper. Fewer than 1% of trials were not assigned to any of the categories

3.3 Interim discussion

Our results show great variation, and while in line with the broad direction of Yun & Lee (2022), they deviate from the canonical view of the relation between prosody and meaning in Korean. Why might this be? There are of course individual differences between experimental participants, but the presence of the observed effects despite the large number of participants suggests that there is something more happening than just a noisy system. Our experiment was large-scale, which may enable patterns to be seen that were not detectable in smaller samples. We also accessed participants online rather than in a controlled laboratory environment: while this can build in variation, it may also more closely reflect the conditions under which language is used day-to-day.

One possible explanation is that there is a lexical bias towards the question-word reading for the relevant CPFs. This is difficult to determine from corpus studies because of the ambiguity between readings, but there is some supporting evidence. For example, the dic.daum.net online dictionary lists the question-word meaning before the indefinite pro-form for all of the words used in our experiment.[4] If this is the case, then it may trigger an early preference among hearers, which may be imperfectly cancelled by the presence of raised levels of F0 at the verb for those polar question stimuli with 100% natural prosody, and which cannot be cancelled for the manipulated stimuli with lower levels of prosody at the verb. This bias may also be present for only some hearers, rather than universally.

[4] Checked 2024-01-24.

4 Challenges to existing LFG work

These results show an ambiguity that all accounts must address, in that the same measurable input (text plus F0 contour) can be interpreted categorically differently by different hearers. In the case of our data, an F0 peak above declination[5] at the CPF is sufficient for interpretation as a content question, but not necessary. Conversely, an F0 peak above declination at the verb is necessary, but not sufficient, for a polar question reading.

Specifically, our results challenge the account proposed by Jones (2016). His specification assumed a definition of a prosodic feature EXPANDED PITCH RANGE that minimally was present at the right edge of the focused constituent and could optionally spread left through the constituent, but which did not appear to the right of the focused constituent. However, for content questions with natural prosody, the right edge of an AccP is not always aligned with the right edge of the focused syntactic constituent. Early attempts to manipulate stimuli such that F0 expansion ended at the right edge of the syntactic constituent were rejected as unnatural by hearers. Acceptable F0 manipulations created a contour with an F0 peak within the AccP with smooth slopes to the AccP boundaries at either side. The presence of a syllable-linked F0 peak is in line with Yun & Lee (2022) and formally has similarities with Mycock & Lowe's (2013) approach to focus linked to English nuclear pitch accent, despite there being scant evidence for a nuclear pitch accent in Korean outside focused contexts.

Jones's account also assumed that the prosodic contribution to the analysis was similar whether EXPANDED PITCH RANGE was associated with the CPF or the verb. The differential conditions for content and polar questions with respect to the relationship between prosodic expression and hearers' interpretation of the string suggest a differential analysis. If CPFs do indeed have a lexical bias towards a content question reading, as discussed in Section 3.3, a satisfactory account will incorporate this.

5 An LFG account

Modelling participants' responses is a conundrum: an optional question-word reading for the CPF must be available, but this must be constrained to prevent overgeneration. The proposed account still assumes the prosodic feature EXPANDED PITCH RANGE proposed by Jones (2016) but this is now defined as being expressed at the F0 maximum of an AccP, rather than minimally at its right edge. There are three new elements: a revised c-structure rule for question focus that makes an explicit link between the presence of a syntactic exponent of question focus and the discourse function FOCUS at i-structure; a revised p-structure rule for the EXPANDED PITCH RANGE feature that associates it with the non-initial H tone in the relevant AccP; and a lexical specification with two alternative entries for indefinite CPFs, the choice of which is constrained at f-structure.[6]

[5]Declination is the tendency for the height of F0 peaks to reduce over the course of an utterance in the absence of other prosodic marking (Ladd 2008).

[6]This proposal using f-structure features is one possible solution: an analysis using p-string elements may provide a more elegant solution and is the subject of further research.

5.1 C-structure rules

Phrase structure rules (2)–(4) are adapted from Cho & Sells (1995) who assume no maximal phrasal projections. In Cho & Sells' rules, the combinatorial possibilities of words are constrained by a feature TYPE, whose value is determined by the particles that attach to the root. Words of any category with TYPE: V-SIS can attach as the left sister of a verb, and words of any category with TYPE: N-SIS can attach as the left sister of a noun. Thus the verb *mekta* 'eat', whose root is *mek*, can host the past tense complementiser *un* with TYPE:N-SIS to modify a noun, e.g. *mek-un sakwa* 'an apple that was eaten', or the tenseless complementiser *e* with TYPE:V-SIS to modify a verb, e.g. *mek-e bota* 'to try to eat'.

Cho & Sells' rules assume only one level of projection in Korean phrases and because they see no evidence for a specifier in Korean syntax, they denote the maximal projection as X′. They take V′ to be the root node of c-structure, but for the purposes of this analysis, which includes cascade rules, we are assuming a root node S. In the full analysis, the type constraints have been omitted for clarity.

(2)
$$\text{S} \rightarrow \begin{array}{c} \text{X}' \\ \text{TYPE: V-SIS} \\ (\uparrow \text{GF}) = \downarrow \end{array} \quad \begin{array}{c} \text{V}' \\ \text{TYPE: NO} \\ \uparrow = \downarrow \end{array}$$

(3)
$$\begin{array}{c} \text{N}' \\ \text{TYPE: Y-SIS} \\ \uparrow = \downarrow \end{array} \rightarrow \begin{array}{c} \text{X}' \\ \text{TYPE: N-SIS} \\ (\uparrow \text{GF}) = \downarrow \end{array} \left\{ \begin{array}{c|c} \text{N}' & \text{N} \\ \text{TYPE: Y-SIS} & \text{TYPE: Y-SIS} \\ \uparrow = \downarrow & \uparrow = \downarrow \end{array} \right\}$$

(4)
$$\begin{array}{c} \text{V}' \\ \text{TYPE: Y-SIS} \\ \uparrow = \downarrow \end{array} \rightarrow \begin{array}{c} \text{X}' \\ \text{TYPE: V-SIS} \\ (\uparrow \text{GF}) = \downarrow \end{array} \left\{ \begin{array}{c|c} \text{V}' & \text{V} \\ \text{TYPE: Y-SIS} & \text{TYPE: Y-SIS} \\ \uparrow = \downarrow & \uparrow = \downarrow \end{array} \right\}$$

5.2 P-structure rules

The p-structure rules (5)–(8) are repeated from Jones (2016) and originally derived from Jun (2005). The rule for prosodic words has been omitted because in Jun's account they play no role in marking phrase boundaries.

(5) Timing tier: p-structure
  a. IntP → AccP^{+}
  b. AccP → Syll^{+}

(6) Timing tier: final syllable lengthening
  Syll → Syll: / ___#

(7) Intonation tier: edge tones
  a. IntP → ___%##
  b. AccP → #TH___LH#

(8) Intonation tier: assimilation of IntP final tones
  H → Ø/ ___%##

5.3 Question semantics

The c-structure and p-structure rules for question semantics are as follows. For the purposes of this paper, we assume that an umbrella abstract notion of 'question semantics' can represent both content and polar question semantics: developing a full glue account that builds on the contributions from Mycock (2006) for content questions and Dalrymple et al. (2019: 422) for polar questions is left for future work. Rule (9) amends rule (16) from Jones (2016), stating that, if present, the syntactic exponent of question semantics **Sem_Qsem** appears within the value of R in the rightmost s-string element of the utterance; the feature CLTYPE with value INT is added to f-structure, indicating an interrogative clause; and the i-structure $\uparrow_{\sigma\iota}$ must have a value for its FOCUS attribute. An S constituent where this holds will include the abbreviated meaning constructor **Qsem** in its s-structure. Rule (10) ensures that the prosodic exponent of question semantics is within the value of the R feature of the rightmost AccP in an IntP, where that IntP has the boundary tune LH%.

(9)
$$\text{S} \rightarrow \text{X}' \quad \begin{array}{c} \text{V}' \\ \left(\begin{array}{c} \textbf{Sem\_Qsem} \in (\searrow \text{R}) \\ (\uparrow \text{CLTYPE}) = \text{INT} \\ (\uparrow_{\sigma\iota} \text{FOCUS}) \end{array} \right) \end{array}$$

(10)
$$\text{IntP} \rightarrow \text{AccP*} \quad \begin{array}{c} \text{AccP} \\ (\% = \text{TONE LH}) \Rightarrow \\ \textit{Sem\_QSem} \in (\searrow \text{R}) \end{array}$$

5.4 Question focus

Syntactic and prosodic rules for question focus, again following Jones (2016) and Mycock & Lowe (2013), are shown below. Rule (11) states that focus is present within a c-structure constituent where **DF_Focus**, the syntactic exponent of focus, is present within the value of the R attribute of its rightmost s-string element. It introduces the f-structure checking feature FOCUS-PROSODY with value +, which is used to constrain the choice of lexical entry for the CPF, discussed further in Section 5.5.

Rule (12) ensures that the s-structure information cascades along the right edge of the constituent headed by the maximal projection of the terminal node associated with the s-string element.

(11)
$$\Sigma \rightarrow \Sigma^* \quad \begin{array}{c} \Sigma \\ \left(\begin{array}{c} (\uparrow \text{FOCUS-PROSODY}) = + \\ (\uparrow_{\sigma} \text{DF}) = \text{FOCUS} \\ \textbf{DF\_Focus} \in (\searrow \text{R}) \end{array} \right) \end{array} \quad \Sigma^*$$

(12)
$$\Sigma \rightarrow \begin{array}{c} \Sigma^* \\ \left\{ \begin{array}{c|c} (\uparrow_{\sigma} \text{DF}) \neq \text{FOCUS} & \begin{array}{c} (\uparrow_{\sigma} \text{DF}) =_c \text{FOCUS} \\ (\uparrow_{\sigma} \text{DF}) = \text{FOCUS} \end{array} \end{array} \right\} \end{array}$$

Rule (13) revises rule (19) from Jones (2016). It specifies that expanded pitch range on the H pitch peak of the AccP is associated with the prosodic exponent of focus. The metavariable $\wr$ in ($\wr$ R) indicates that the prosodic exponent is associated with feature R of the syllable that it relates to, not to the edge of the containing AccP. Following Jones (2016), Interface Harmony in Korean can be satisfied by any of the p-string elements corresponding to an s-string element: the relevant p-string element does not need to be at the edge of a p-structure constituent.

(13) $$\text{AccP} \rightarrow \text{Syll*} \quad \underset{\substack{\big[(\text{TONE} = \text{H}) \wedge \\ (\text{PITCH} = \text{EXP})\big] \Rightarrow \\ \textit{DF\_Focus} \in (\wr\ \text{R})}}{\text{Syll}} \quad \text{Syll*}$$

5.5 A lexical bias towards content question readings

The lexical bias of CPFs towards content question readings is modelled by the two alternative lexical specifications in (14), one with a baseline reading which assumes that focus is assigned prosodically, and one which assigns question focus to the CPF in the absence of prosodic support.

The baseline specification (14a) is always available. In the version that assigns question focus to the CPF, (14b), there are three additional constraints. In the f-description, the constraint $((\text{OBJ} \uparrow)\ \text{CLTYPE}) =_c \text{INT}$ prevents the question word reading from appearing in non-question utterances. In the p-form, the constraint $\textit{DF\_Focus} \in (\wr\ \text{R})$ adds the feature of prosodic exponence, *DF_Focus*, into the right edge set of the AccP containing the CPF. This allows the exponent feature to appear in the p-form even if the prosodic feature as specified in rule (13) is not present, in turn allowing the constraint in (11) to be satisfied.

(14) a. Baseline lexical entry with ambiguous reading:

s-form	$(\bullet\ \text{FM}) =$	nwukwulul
c-structure category	$\lambda(\pi(\bullet)) =$	N [TYPE:V-SIS]
f-description	$(\uparrow \text{PRED}) =$	'PRO'
	$(\text{OBJ} \uparrow)$	
p-form	/nuguruɯl/	

b. Lexical entry with obligatory focus

s-form	$(\bullet\ \text{FM}) =$	nwukwulul
c-structure category	$\lambda(\pi(\bullet)) =$	N [TYPE:V-SIS]
f-description	$(\uparrow \text{PRED}) =$	'PRO'
	$(\text{OBJ} \uparrow)$	
	$((\text{OBJ} \uparrow)\ \text{CLTYPE}) =_c \text{INT}$	
	$((\text{OBJ} \uparrow)\ \text{FOCUS-PROSODY}) \neq +$	
p-form	/nuguruɯl/	
	$\textit{DF\_Focus} \in (\wr\ \text{R})$	

5.6 The analysis

The three cases—content question correctly identified, polar question correctly identified and polar question misidentified as a content question—can then be analysed using the proposal from Jones (2016), as amended in (9)–(14). All examples use sentence (1), repeated here.

(1′) *hakchangsicel ttay* *nwukwu-lul* *mollay* *sarangh-aysse-yo*
school.days during who/someone-OBJ secretly love-PST-POL

a. “You secretly loved someone when you were at school.” (statement)
b. “Who did you secretly love when you were at school?” (content question)
c. “Did you secretly love someone when you were at school?” (polar question)

For clarity, in Figures 7–9, c- and p-structures for *hakchangsicel ttey* ‘at school’ have been simplified, adjuncts have been omitted from f-structure, and the mapping from s-structure to i-structure is not shown.

5.6.1 Content question

If EXPANDED PITCH RANGE is perceived to be present at the CPF, rule (13) applies to the AccP containing the CPF. This sets the value of f-structure feature FOCUS-PROSODY to be $+$ and so prevents the use of lexical entry (14b). The question is interpreted as a content question because the feature **DF_Focus**, the syntactic exponent of focus, appears in the right edge set of the CPF constituent, satisfying the principle of Interface Harmony. Figure 7 shows the full analysis of sentence (1) with perceived pitch expansion at the CPF, resulting in reading (1b).

5.6.2 Polar question

If EXPANDED PITCH RANGE is perceived to be present at the verb, rule (13) applies to the final AccP of the utterance, again setting the value of f-structure feature FOCUS-PROSODY to be $+$ and preventing the use of lexical entry (14b). In this case, the question is interpreted as polar, with the feature **DF_Focus** in the right edge set of the verb. Again, the principle of Interface Harmony is satisfied. Figure 8 shows the full analysis of sentence (1) with perceived pitch expansion at the verb, resulting in reading (1c).

5.6.3 Polar question misidentifed

We assume that misidentification of polar questions occurs because EXPANDED PITCH RANGE is not perceived at the verb. In this case, where there is still a final LH% tune, the utterance is interpreted as a content question. Rule (9) requires the discourse function FOCUS to be present at i-structure, without specifying which element should be in focus. However, rule (13) does not apply: without Interface Harmony the optional focus constraints in rules (11)–(12) cannot be applied. The c-structure rules do not provide a value for $(\uparrow_{\sigma\iota}$ FOCUS$)$ (mediated via s-structure) and the f-structure feature FOCUS-PROSODY is unspecified. In order to satisfy the requirement for focus from rule (9), lexical entry (14b) must be used, which in turn assigns question focus to the CPF using Interface Harmony. Figure 9 shows the full analysis of sentence (1) with no perceived pitch expansion, resulting in reading (1b) rather than the speaker’s intention (1c).

6 Discussion

The rules above can generate a well-formed analysis whether or not EXPANDED PITCH RANGE is perceived at either the CPF or the verb. However, they do not explain why the variation in perception occurs. This paper does not seek to explore the wider issue of divergent speaker and hearer analyses: in our view it is sufficient to ensure that there are reasonable formal assumptions about the nature of the speaker and hearer experience that support a well-formed account. However the question remains why a substantial proportion of hearers perceived a content question even though the natural prosody associated with the production of a polar question was in the signal. The difference between our results and the canonical view merits further discussion and potentially further experimental exploration.

In the interim discussion, Section 3.3, the possibility was mentioned of an "early prediction, imperfect cancellation" strategy, linked to a lexical preference for the question-word reading of the CPF. If that is the case, then the gating paradigm, with its repetition of the earlier stimuli, may have reinforced the prediction. It seems that most participants understood the instructions as being only to give a judgement when they were certain what the type of utterance was, leading to a delay in identification. Most participants answered "don't know" until segment 4 had been heard. However, from segment 2, when the CPF was heard, through to segment 4, the verb, the second most frequent answer, for all proportions of natural prosodic range, was to identify the stimulus as a content question—see Jones et al. (under review) for more details. This provides tentative support for the idea of an imperfectly-cancelled prediction.

The algorithm used to manipulate the stimuli may also have introduced a confound. In a pilot attempt where the pitch was manipulated at many intermediate points, the manipulated stimuli were unintelligible and so we were advised to take a simpler approach. However, it is possible that the F0 contours, although tested for acceptability with native speakers, were oversimplified and that F0 contours are needed with more intermediate scaling points.

Further experimentation using refined stimuli could help to build an explanation, with two modes of stimulus presentation—the gating paradigm and entire utterances—and revised instructions asking participants to say as early as possible which type of utterance they think is the most likely, depending on what they have heard. If it appears that participants change their opinion during the course of the utterance (other than from "don't know" to one of the sentence types) the phenomenon would be amenable to incremental cognitive modelling.

7 Conclusion

We entered this project with the aim of sharpening our understanding of the relationship between F0 contours and meaning in otherwise ambiguous Korean sentences. Our data suggest not only that this relationship is different to and more complex than previous accounts would have it, but also that the canonical assumption of the central role of prosody in disambiguation is mediated by lexical factors. This merits further investigation. Our proposed analysis uses LFG's established techniques for dealing with ambiguity and allows perceptual differences between hearers to be included in a unified account.

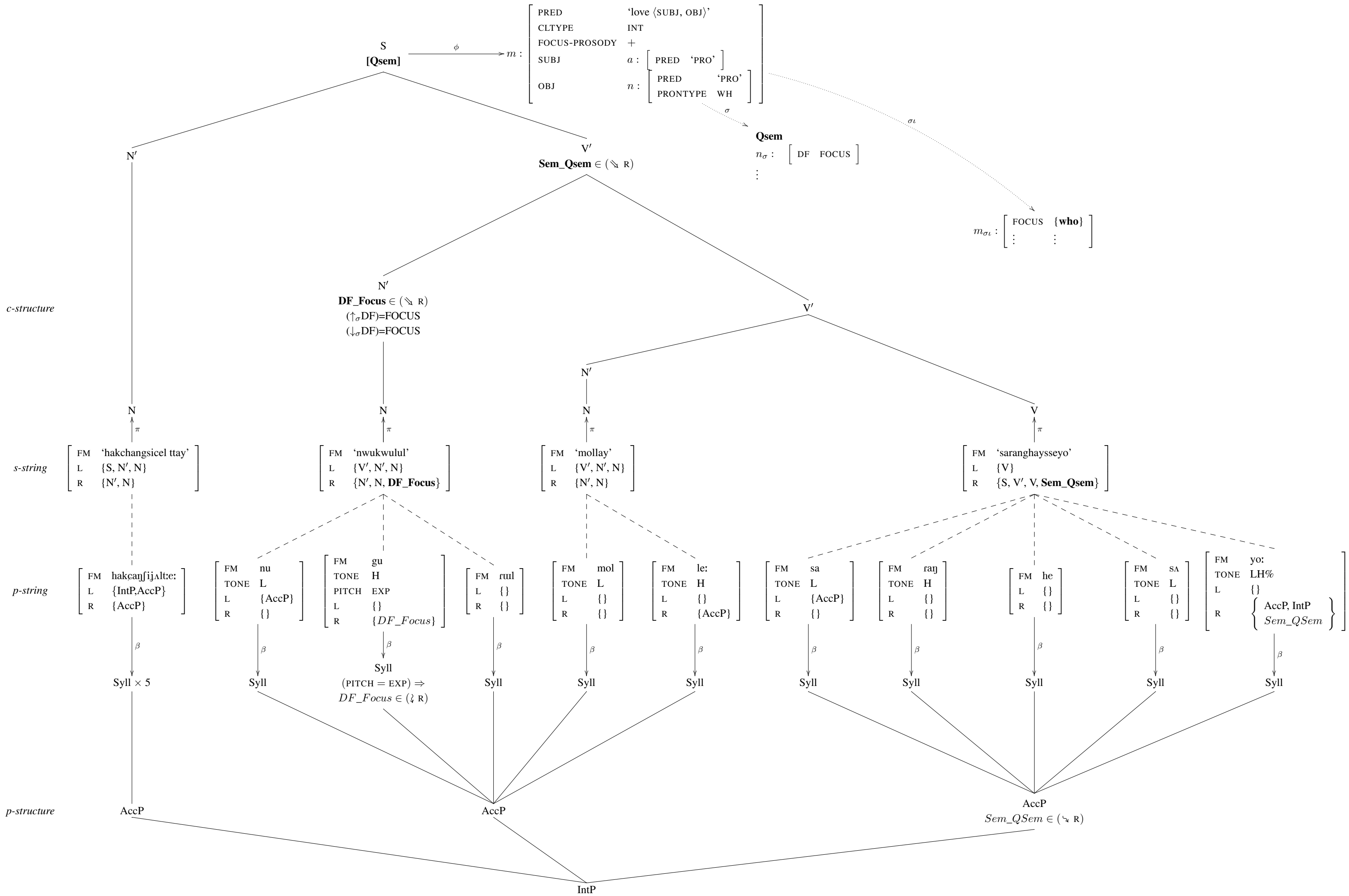

Figure 7: Content question: 'Who did you secretly love at school?'
Prosodic feature EXPANDED PITCH RANGE is perceived at syllable [gu] during the CPF *nwukwulul* 'who/someone', placing the prosodic exponent *DF_Focus* in the value of its R feature by rule (13). Interface Harmony requires the syntactic exponent **DF_Focus** in the value of the R feature of the s-string element 'nwukwulul'. Note that the right edges of N 'nwukwulu' and AccP [nuguɾɯl molleː] do not align

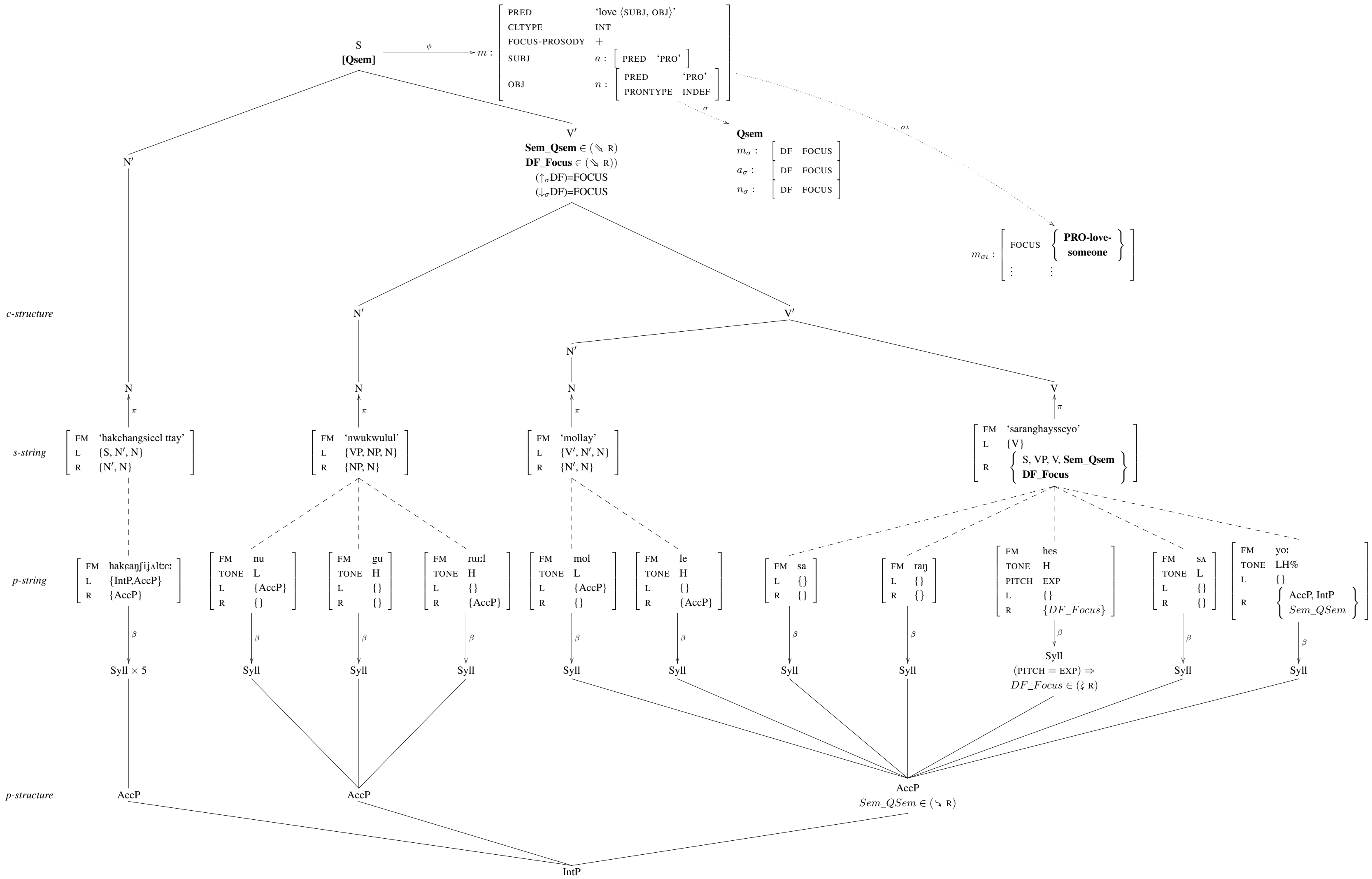

Figure 8: Polar question: 'Was there someone that you secretly loved at school?'
Prosodic feature EXPANDED PITCH RANGE is perceived at syllable [hes] during the verb *saranghaysseyo* 'who/someone', placing the prosodic exponent *DF_Focus* in the value of its R feature by rule (13). Interface Harmony requires the syntactic exponent **DF_Focus** in the value of the R feature of the s-string element 'saranghaysseyo'

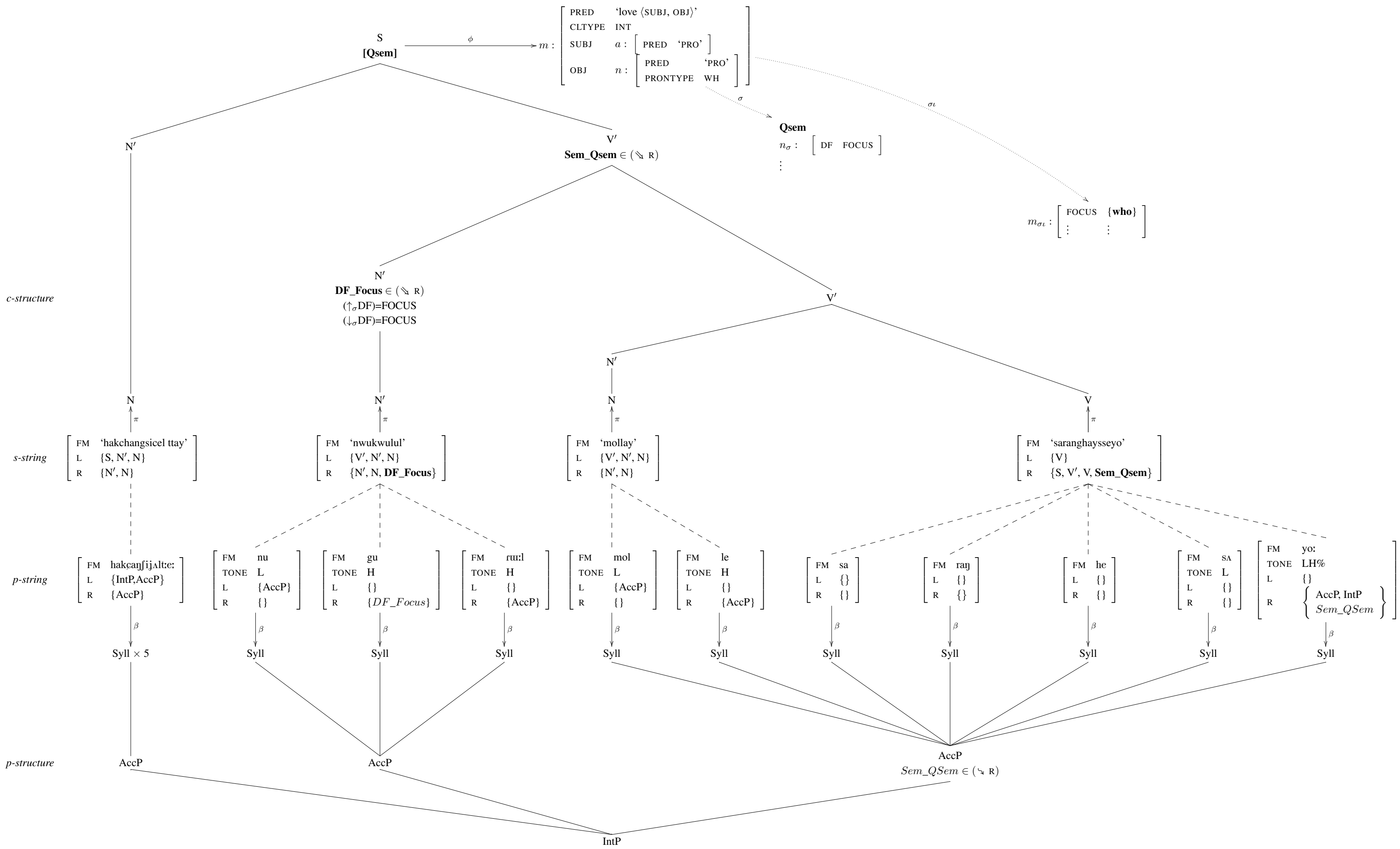

Figure 9: Polar question 'Was there someone that you secretly loved at school?' where EXPANDED PITCH RANGE during the verb is not perceived by the hearer so the question is incorrectly identified as a content question

References

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01.

Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5(9/10). 341–345.

Bögel, Tina. 2015. *The syntax-prosody interface in Lexical Functional Grammar*. Ph.D. thesis, University of Konstanz.

Bögel, Tina. 2022. The prosody-syntax interface: A computational implementation. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG'22 conference*, 61–78. Stanford, CA: CSLI Publications. https://lfg-proceedings.org/lfg/index.php/main/article/view/1/8.

Bögel, Tina. 2023. Prosody and its interfaces. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*, 779–821. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.10185969.

Cho, Young-Mee Yu & Peter Sells. 1995. A lexical account of inflectional suffixes in Korean. *Journal of East Asian Linguistics* 4(2). 119–174. https://doi.org/10.1007/BF01731614.

Dalrymple, Mary, John J. Lowe & Louise Mycock. 2019. *The Oxford reference guide to Lexical Functional Grammar*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198733300.001.0001.

Dalrymple, Mary & Louise Mycock. 2011. The prosody-semantics interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'11 Conference*, 173–193. Stanford, CA: CSLI Publications. https://typo.uni-konstanz.de/lfg-proceedings/LFGprocCSLI/LFG2011/papers/lfg11dalrymplemycock.pdf.

Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511993473.

Hastie, Trevor J. & Robert J. Tibshirani. 1990. *Generalized additive models*. London: Chapman & Hall/CRC.

Jones, Stephen M. 2016. The syntax-prosody interface in Korean: Resolving ambiguity in questions. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 318–338. Stanford, CA: CSLI Publications. https://doi.org/10.21248/hpsg.2016.17.

Jones, Stephen M., Yoolim Kim & Cong Zhang. Under review. Perceiving and modelling the scope of question focus in Korean.

Jun, Sun-Ah. 2005. Korean intonational phonology and prosodic transcription. In Sun-Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*, 202–228. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199249633.003.0008.

Jun, Sun-Ah & Mira Oh. 1996. A prosodic analysis of three types of wh-phrases in Korean. *Language and Speech* 39(1). 37–61. https://doi.org/10.1177/002383099603900103.

Ladd, D. Robert. 2008. *Intonational phonology* (2nd edn.). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO97805118088140.

Lange, Kristian, Simone Kühn & Elisa Filevich. 2015. "Just Another Tool for Online Studies" (JATOS): An easy solution for setup and management of web servers supporting online studies. *PLoS ONE* 10(7). e0130834. https://doi.org/10.1371/journal.pone.0130834.

de Leeuw, Joshua R., Rebecca A. Gilbert & Björn. Luchterhandt. 2023. jsPsych: Enabling an open-source collaborative ecosystem of behavioral experiments. *Journal of Open Source Software* 8(85). 5351. https://doi.org/10.21105/joss.05351.

Lennes, Mietta. 2017. The speech corpus toolkit for Praat. https://doi.org/10.5281/zenodo.375923.

Mycock, Louise. 2006. *The typology of constituent questions: A Lexical-Functional Grammar analysis of 'wh'-questions*. Ph.D. thesis, University of Manchester.

Mycock, Louise & John Lowe. 2013. The prosodic marking of discourse functions. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG13 Conference*, 440–460. Stanford, CA: CSLI Publications. https://typo.uni-konstanz.de/lfg-proceedings/LFGprocCSLI/LFG2013/papers/lfg13mycocklowe.pdf.

R Core Team. 2023. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org.

Selkirk, Elizabeth O. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

van Rij, Jacolien, Martijn Wieling, R. Harald Baayen & Hedderik van Rijn. 2022. itsadug: Interpreting time series and autocorrelated data using GAMMs. R package version 2.4.1. https://cran.r-project.org/web/packages/itsadug/.

Wood, Simon N. 2017. *Generalized additive models: An introduction with R* (2nd edn.). New York: Chapman and Hall/CRC. https://doi.org/10.1201/9781315370279.

Yun, Jiwon. 2019. Meaning and prosody of *wh*-indeterminates in Korean. *Linguistic Inquiry* 50(3). 630–647. https://doi.org/10.1162/ling_a_00318.

Yun, Jiwon & Hye-Sook Lee. 2022. Prosodic disambiguation of questions in Korean: Theory and processing. *Korean Linguistics* 18(1). 18–47. https://doi.org/10.1075/kl.00014.yun.