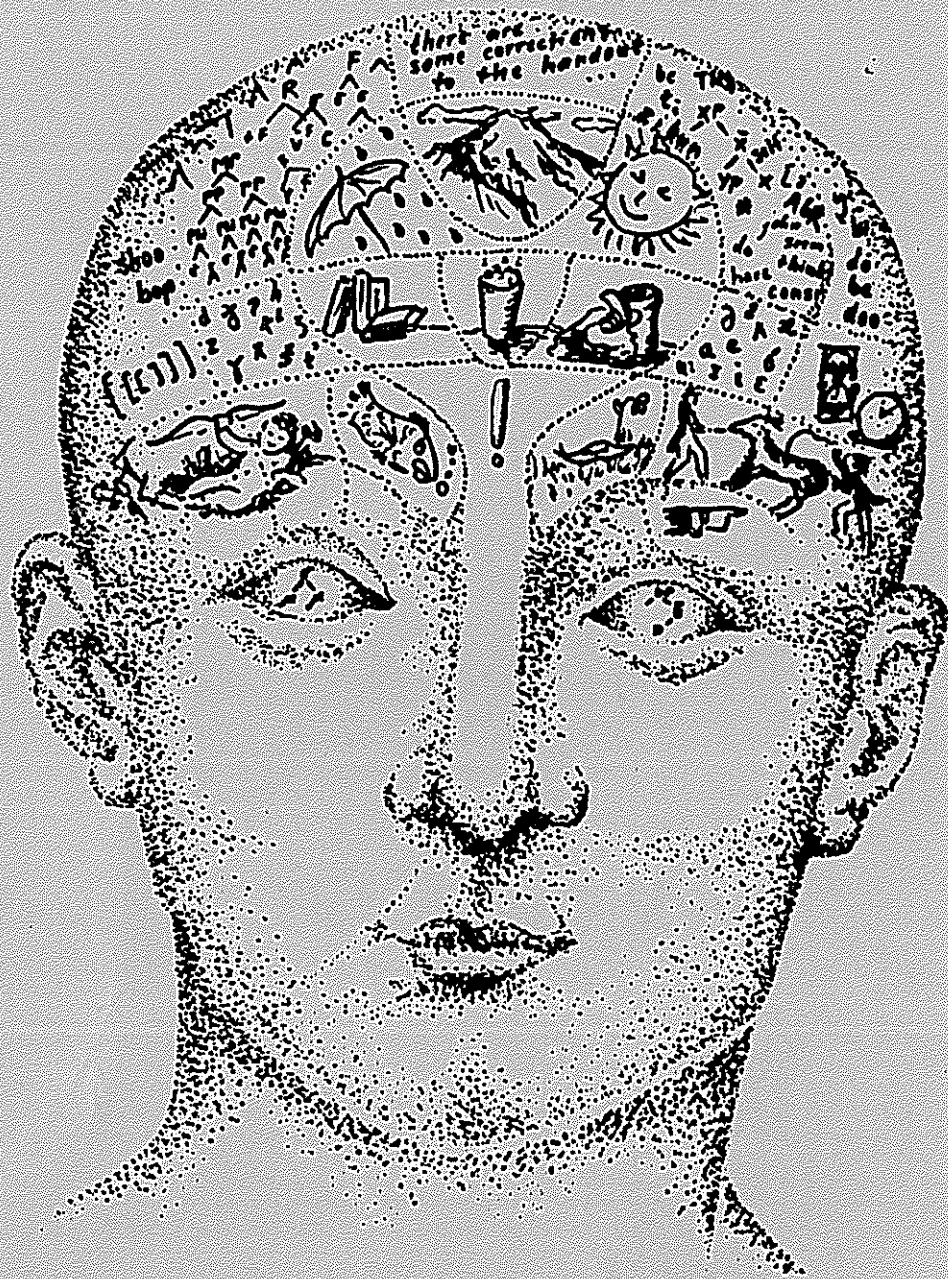


WORKING PAPERS
IN LINGUISTICS
UNIVERSITY OF WASHINGTON



VOLUME 13
JULY 1996

Errata

Working Papers in Linguistics
volume 13, July 1996

Inside front cover

Tatsuya Suzuki (Ph.D. Linguistics 1987)

should read

Tatsuya Suzuki (Ph.D. Linguistics 1988)

Table of Contents (also on back cover)

Acoustic Correlates of Dec Xinag Vowels

should read

Acoustic Correlates of Deg Xinag Vowels

Page 1, footnote 1

...future progressive...

should read

...future **and** progressive...

Inside back cover

Kyu Hwang

should read

Kyu-**Hong** Hwang

Working Papers in Linguistics
University of Washington
Volume 13, July 1996

Editor: Hideo Makihara

Editorial Assistant:
James Lyle

Special Thanks to: Tatsuya Suzuki (Ph.D. Linguistics 1987)

For subscription, exchange, and advertising information, contact the editor at

Department of Linguistics
University of Washington
Box 354340
Seattle WA 98195-4340 USA
(206) 634- 2046
phoneme@u.washington.edu

All papers in this volume remain the sole property of the author and may be protected by copyright; permission to duplicate should sought from the individual author.

Cover graphics: Siri Tuttle

WORKING PAPERS IN
LINGUISTICS
UNIVERSITY OF WASHINGTON

VOLUME 13
JULY 1996

Department of Linguistics
University of Washington
Box 354340
Seattle, WA 98195-4340
(206) 634-2046
phoneme@u.washington.edu



TABLE OF CONTENTS

Acoustic Correlates of Dec Xinag Vowels Alice Taff	1
Licensing and Scope Interpretation of Wh-Phrases in Korean Tai-Soo Kim	33
A Case-Tier Approach to Finnish: Feature-based Definition of Case Domain Yongkil Jeong	75
The English Subjunctive and Minimalist Case Checking Theory Kyu-Hong Hwang	95
The Evolution of Pronominal <i>ille</i> from Latin to Spanish: A Grammaticalization Analysis José Carrasquel	121
The 'double-go-future' construction in Spanish: A New Grammaticalization phase? Inés Miranda	147

ACOUSTIC CORRELATES OF DEG XINAG VOWELS*

Alice Taff

1995

1. INTRODUCTION

Objective The objective of this study is to determine the acoustic correlates of the vowel in the second person plural morpheme in Deg Xinag (Ingalik Athabaskan) and compare these vowel correlates to those of other vowels in the Deg Xinag vowel inventory. The identity of the vowel in the second person plural morpheme is controversial among linguists, transcribed as either /oχ/ (Hargus and Taff 1991) or /uχ/ (Kari 1991).¹ Orthographic renderings by native speakers do not set the controversy to rest as there is relatively little writing taking place in this language. Deg Xinag speaker/writers typically verify spellings with dictionaries and word lists generated by linguists or have linguists proofread texts to be published.²

Hereafter, I will refer to the second person plural morpheme vowel as V2p. The question to be addressed here is: can acoustic measurements of this vowel contribute to its phonemic disambiguation? If the acoustic correlates of the second person plural vowel pattern like those of another Deg Xinag vowel in a similar environment, then it is that vowel. If its correlates pattern differently than other Deg Xinag vowels, then it may be a completely different phoneme or may be one of the established vowels undergoing a shift.

Part 2 gives methods of procedure. Part 3 gives results of the measurements. Part 4 discusses the results. In Part 4.1, I draw the conclusion that, although V2p shares acoustic correlates with /ə/, /o/, and /u/, it is best spelled as "u". In addition, the results of these vowel measurements of Deg Xinag give rise to discussion in Part 4.2 of the non-use of high vowel space in this language.

* This study was supported by Alaska Humanities Forum grant #47-90, the University of Washington, and National Science Foundation Instrumentation and Laboratory Improvement Program grant DUE-9352231. My thanks to those who offered suggestions on this paper, Kate Davis, Sharon Hargus, Ellen Kaisse, Carol Stoel-Gamon, Jim Kari, and Siri Tuttle.

¹Kari predicts /oχ/ in future progressive forms only (Jim Kari 1991, p.c.).

² There is a related problem in writing the vowel of the 'areal' prefix, the prefix that refers, sometimes very abstractly, to area. The areal prefix is written variously as *xi* /χə/ and *xu* /χu/. The word 'story' has been published as *xudhoyh* and *xidhoyh*. The word 'peoplé' has been published as both *xut'an*, and *xit'an*. (Jim Kari 1996, p.c.)

Background With speakers on the western Yukon and Innoko (and formerly Kuskokwim) Rivers in Alaska, Deg Xinag is the westernmost language in the Athabaskan language family, which itself extends eastward through central Alaska and Canada, then southward, discontinuously, through the U.S. to Navajo and Apache territory. At this time, there are between twenty and thirty Deg Xinag speakers, depending on fluency definitions.

Descriptive accounts (Kari 1978, Leer 1979, Hargus and Taff 1991) document for Deg Xinag, as for many Athabaskan languages (Krauss and Golla 1981), a relatively small vowel inventory of four phonemes, /e, a, o, ə/ (Krauss and Golla 1981), or five phonemes: three "long" (full) vowels /e, a, o/ and two "short" (reduced) vowels /ə, u/ (Kari 1978, Leer 1979, Hargus and Taff 1991).³ In either case, the vowel system is an unusual one in that it lacks /i/ and /u/, found in virtually all other languages (Crothers 1978, Maddieson 1984). Deg Xinag vowel trapezoids are given in Figures 3-6 and 14-17.

Some Deg Xinag /u/'s are the reflex of the labial component of a Pre Proto-Athabaskan labialized uvular series, Q^w (Leer 1979:15). /u/ is restricted in Deg Xinag to pre and post uvular environments, i.e., /uQ/ or /Qu/.

The Deg Xinag phonemic consonant inventory is given in Table 1.

³ Minimal pair sets indicate that the five vowel system is correct:

e - ə	/ðedo/ 'you're sitting'	/ðədo/ 'he is sitting'
ə - u	/xəy/ 'winter'	/xuy/ 'roots'
u - o	/q'uθ/ 'cloud'	/q'oθ/ 'quiver; exercise'
o - a	/voxo/ 'for him or her on his or her behalf'	/vaça/ 'with it'

Table 1. Deg Xinag phonemic consonant inventory. ' indicates [+ glottal]. ˘ indicates [-voice].

tongue	coronal						dorsal		
other articulator	labial	dental	alveolar			palatal	velar	uvular	glottal
manner				lateral	retroflex				
stops		d					g	ɢ	
	p	t					k	q	
		t'					k'	q'	ʔ
affricates		dð	dz	dl	dr	ʃ			
		tθ	c	tl	tr	č			
		tθ'	c'	tl'	tr'	č'			
fricatives	v	ð	z		ʒr			ʁ	
		θ	s		sr	ʃ		χ	h
sonorants	m	n		l		y	ŋ		
		n̥		l̥		y̥	ŋ̥		
	m'	n'				y'	ŋ'		

To my knowledge, while there has been documentation and description of the phoneme inventories of all the Athabaskan languages (e.g. Krauss 1964, Krauss and Golla 1981), the acoustic studies of these systems are few (but see Latimer 1978, McDonough, et al., 1993, and McDonough and Ladefoged 1993).

2. METHODS OF PROCEDURE

Data were derived from recordings made for Jerue 1993, Hargus, et al in prep, and Hargus and Taff 1991. The native Deg Xinag speakers whose voices were recorded and measured for this project are two women in their sixties. Deg Xinag is their first language and both are now fluent speakers of English as well. One is also fluent in Holikachuk, the Athabaskan language spoken directly north of Deg Xinag territory on the Yukon and Innoko Rivers. During the audiotape recording sessions in September 1991, the consultants were seated at a table in a room with two researchers. Utterances were recorded using a Sony Walkman Professional tape recorder, model WM-D6 and Sony electret condenser microphone model ECM-1929LT at 1 7/8 ips.

For these recordings, speakers gave citation forms of minimal and near-minimal pairs with crucial sound contrasts. Stimuli were oral English

translations provided by the researcher as well as written English and DX forms. One speaker or the other produced the first example which was immediately repeated by the second speaker. With the same kinds of stimuli, the speakers gave verb paradigms which included the second person plural form, the morpheme with the vowel 2p. Additional data is extracted from conversations and narratives recorded during these sessions.

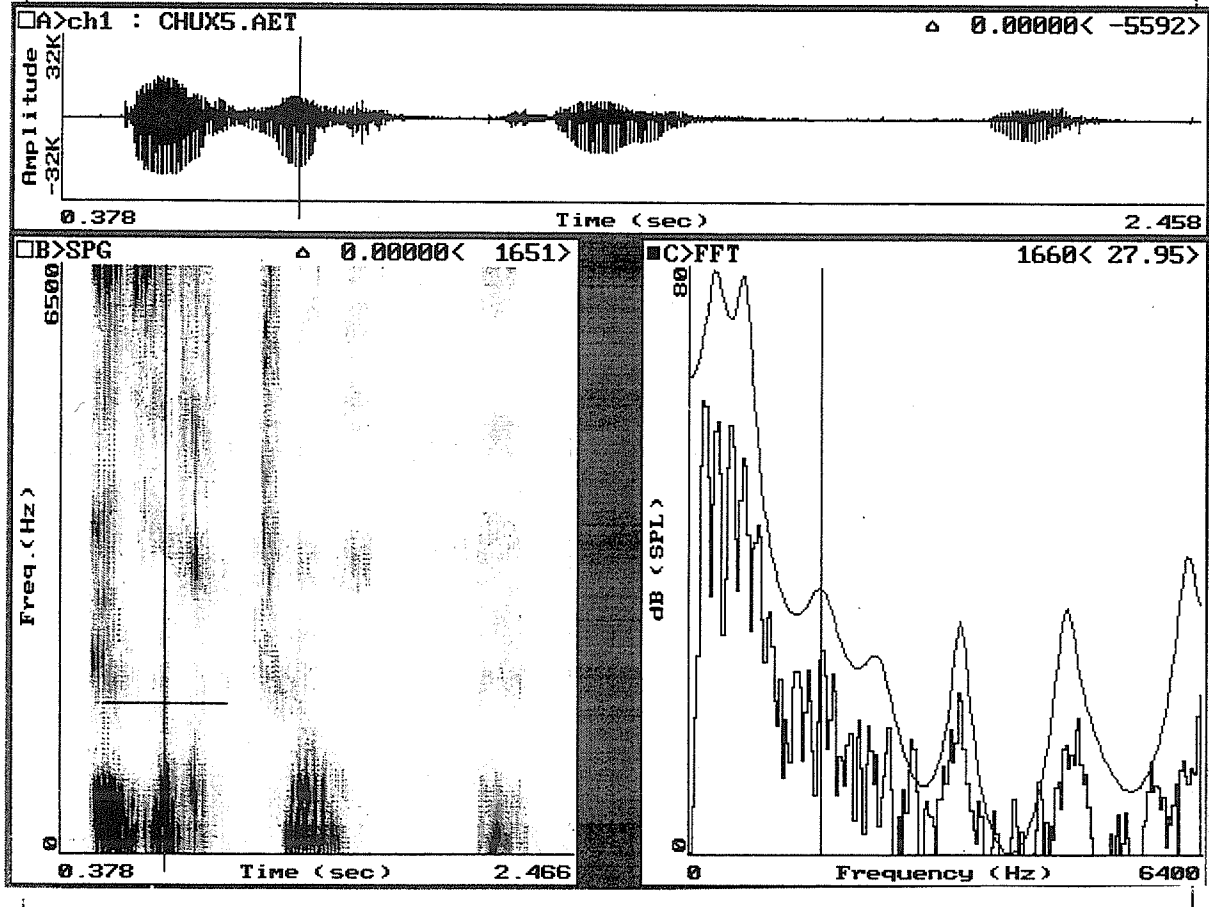
Of these speech types, the citation forms are the slowest and most 'carefully' pronounced. The conversations, intended as examples of conversation for Deg Xinag learners, are assumed to be somewhat slower than they would be in a real world communication situation. Even at that, they are faster than the citation forms. And the narrative speech is the most rapid. Around 90% of the data in this project is taken from the citation forms. Vowel examples taken from the conversation and narrative text were excluded from the duration results but were reported in the formant measurements.

From these data, transcriptions by Hargus, Kari, Jerue, Maillelle, and Taff, who each reviewed the material, were the basis for phoneme assignment for each vowel. That is, if a vowel is transcribed as /o/, it was measured, then analysed here along with the rest of the data transcribed as /o/. Transcriptions are robustly uniform across transcribers, one of the exceptions being the transcription of V2p.

In the sense that Jerue and Maillelle, who both worked on the transcription, are native speakers of Deg Xinag, we can say that perception tests by native speakers have been performed on the data. Jerue et al. (1993) gives the orthography of the second person plural morpheme as both *ux*, phonemically /uχ/, and *ox*, phonemically /oχ/ (e.g., pp 44-46).

Acoustic measurement proceeded as follows. I measured, in four environments, the first three formants for each vowel, /a/, /e/, /u/, /o/, /ə/ and V2p, using a computerized speech lab system from Kay, CSL 4300. The data were digitized at 16,000 kHz. F1, F2, and F3 were measured from spectral analysis (bandwidth 254 Hz) confirmed with a power spectrum as in Figure 1. The utterance is *edhuxchux he'* /eðV2pχčvχ he?/ 'are you all big'. The spectrum is taken from the center of V2p with an F2 reading marked by the cursor in window C where the F2 measurement of 1600 Hertz is given in the upper right hand corner.

Fig. 1. Wave, spectrogram and spectral display of *edhuxchux he'* /eðV2pχçvχ heʔ/ 'are you all big' with F2 spectrum reading of V2p. The formant measurements here are: F1 = 670, F2 = 1660, F3 = 3440.



Linear predictive coding (LPC) was used to generate formant histories for ambiguous cases.

The phonetic environments are:

- coronal V coronal (CVC)
- coronal V uvular (CVQ)
- uvular V coronal (QVC)
- uvular V uvular (QVQ).

Of the vowels in these environments, those preceded or followed by sonorants were excluded from the database on the grounds that sonorants have a damping effect in the higher frequencies.⁴ The database strives towards ten examples of

⁴In addition, data were deleted when there were audible phone calls, footsteps, door slams,

each of the six vowels in each of the four environments. Table 2 gives the number of tokens measured for each vowel, a total of 225 tokens.

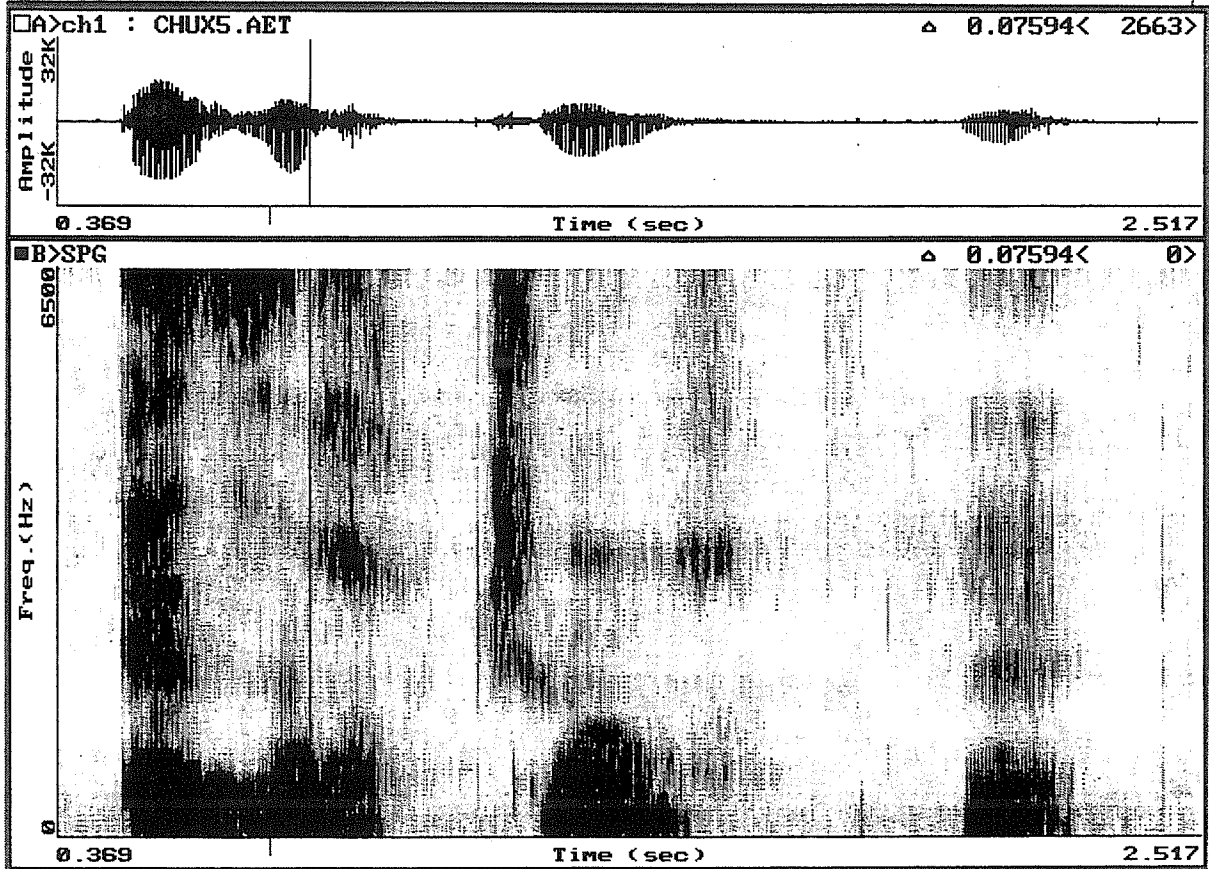
	CVC		QVC		CVQ		QVQ	
	stressed	unstressed	stressed	unstressed	stressed	unstressed	stressed	unstressed
/a/	3	9	6	4	13	0	14	0
/e/	6	4	0	6	8	2	4	0
/o/	8	4	4	0	10	0	5	0
/ə/	7	6	8	5	12	5	9	1
/u/	NA	NA	15	0	35	0	12	0
2p	NA	NA	NA	NA	NA	10	NA	6

Vowels with fewer than the target of ten examples were exhaustively measured in the data available. For the more frequently occurring vowels, more data is available so more measurements were taken. Both stressed and unstressed syllables were measured and stress is accounted for in the analysis.

Duration measurements were made from the spectrograms, using both visible and audible cues. The visual cue for the beginning of duration was the onset of both F0 and formant structure. The ending point was at the formant pattern's complete disintegration or the beginning of the following consonant. I adjusted the measurement of the visual pattern by listening to the portion marked for measurement and deleting from the time any audible preceding or following consonant. See Figure 2 for an example of vowel duration measurement. Here, as in Figure 1, the utterance is *edhuxchux he'* /eðV2pχčvχ he? / 'are you all big'. The 76 ms. duration measurement is of V2p. The cursor marks in both the waveform, window A, and spectrogram, window B, are linked marking the beginning point (below the frame) and ending point (through the frame) for this duration measurement.

snowmachine passings, or sled dog howls in the signal.

Fig. 2. Duration measurement of V2p in *edhuxchux he'* [eðV2pχçuχ he'] 'are you all big'. Cursor markers at its beginning and end of measure V2p at 76 ms.



In McDonough et al. (1994), variables were limited by analysing only stem vowels, all of which receive primary stress. In the research reported here, both stressed and unstressed syllables were measured since the object of the investigation, the second person plural morpheme, is unstressed in these data.

3. RESULTS

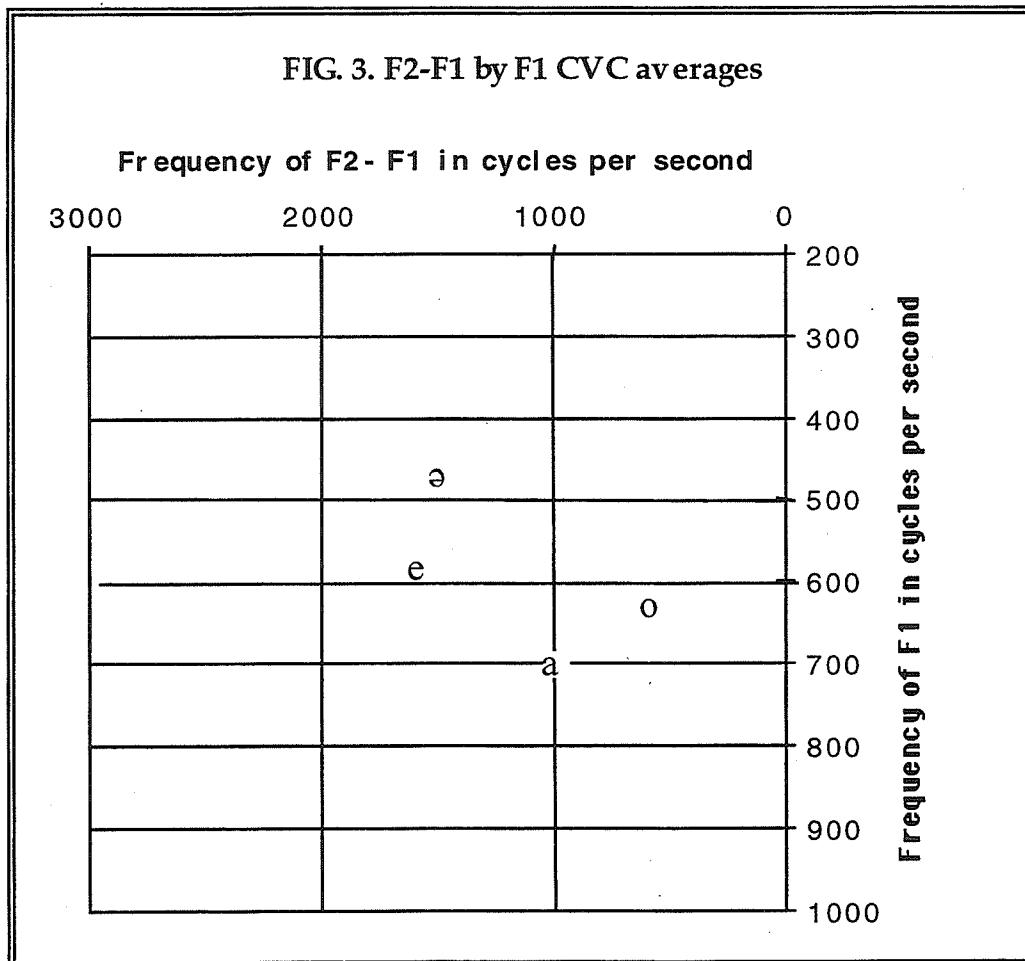
3.1 FORMANT STRUCTURE

In this study, F₁ and F₂ were measured and plotted to show the distribution of the vowels in vowel space. Scatter plots of the data show much overlap among the vowels and are given as Appendix A. Plots of the averages for each

vowel are more illustrative for the purpose of this paper.

Figures 3 through 6 below plot F2-F1 by F1 averages⁵ of each of the vowels in the four environments, CVC - Figure 3, QVC - Figure 4, CVQ - Figure 5, and QVQ - Figure 6. Figures 5 and 6 show the two phonetic environments in which V2p appears, since the second person plural morpheme is V2p followed by /χ/.

Figure 3 gives the positions (as averages) of the vowels when in the environment of a preceding and following coronal consonant. The vowel /u/ does not appear here as it is only transcribed next to a uvular.



⁵ F2-F1 is the correlate for backness, the smaller the difference, the further back the vowel. F1 is the correlate for height, the higher F1, the lower the vowel.

Figure 4 shows evidence of uvular coarticulation as the front vowels are much further back relative to the same vowels shown in Figure 3's purely coronal environment. Figure 4 introduces the vowel /u/.

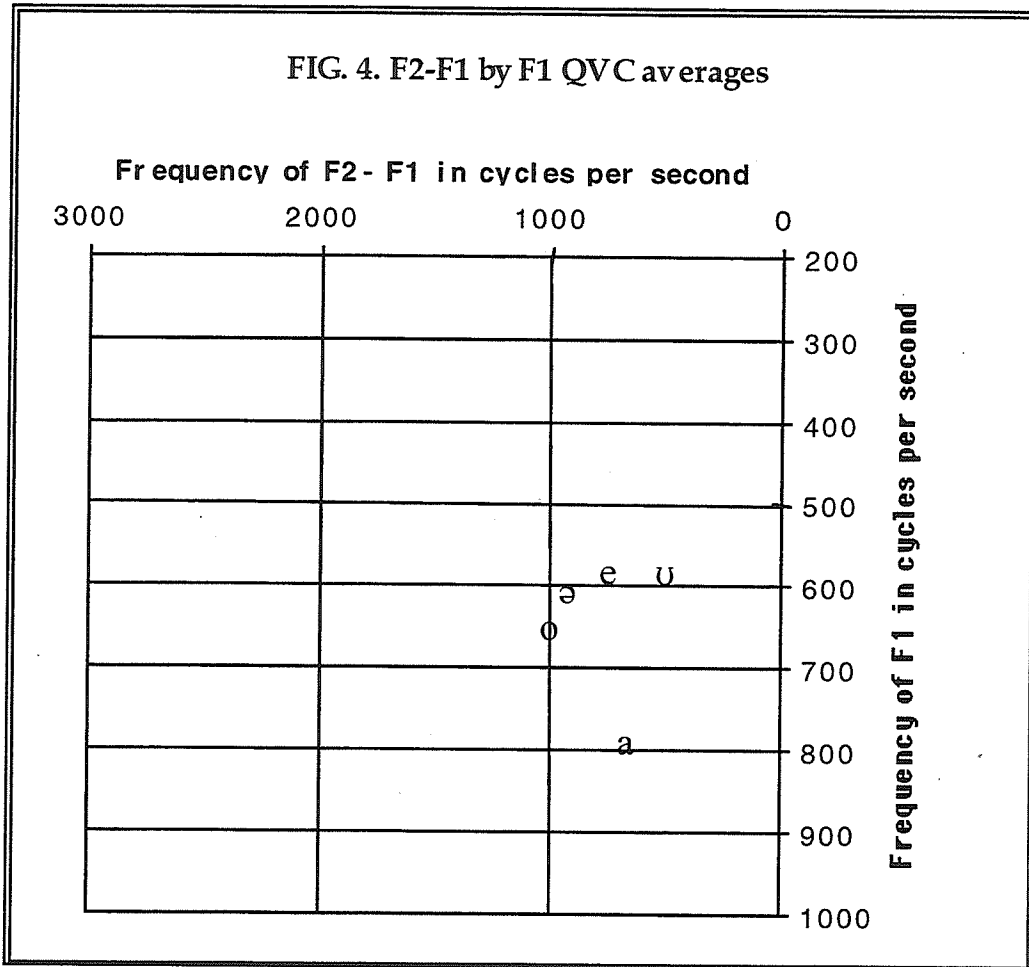


Figure 5 shows that when preceded by a coronal and followed by a uvular, V2p is closest to /u/ and /o/, but these three vowels and /ə/ are very close together.

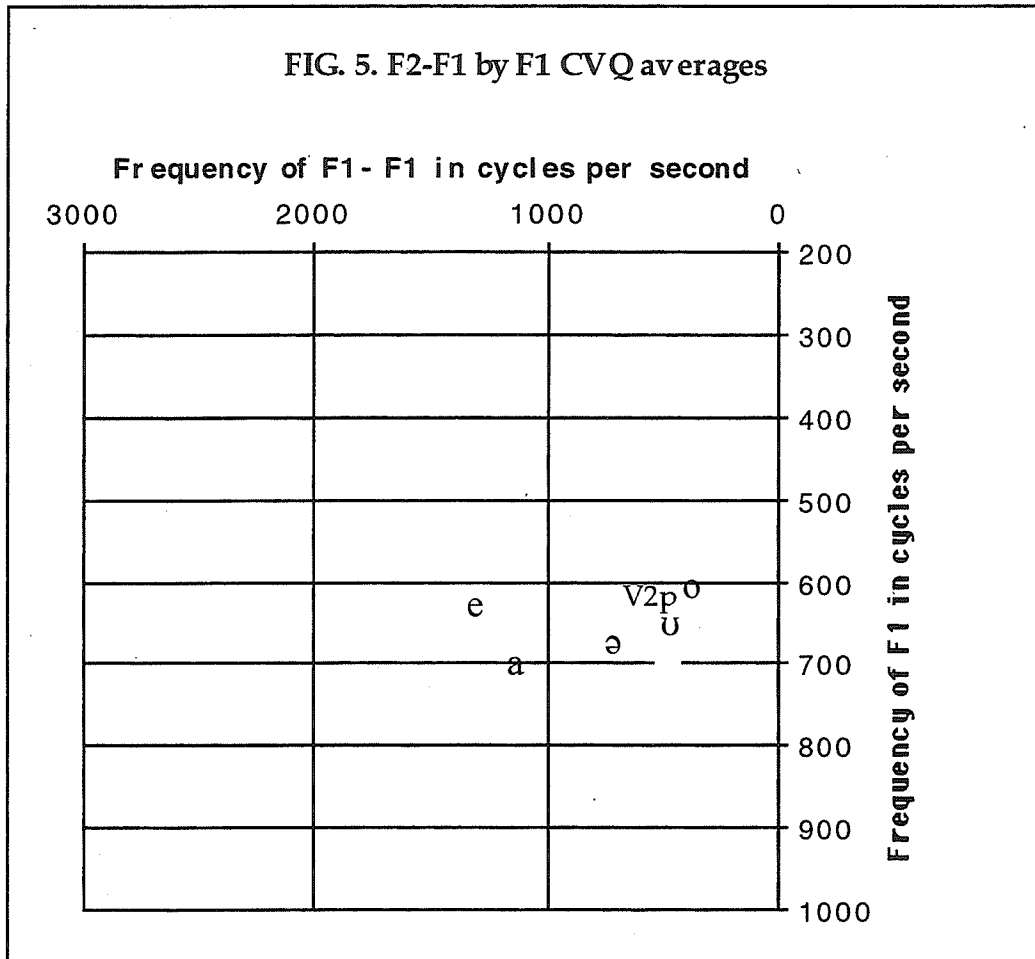
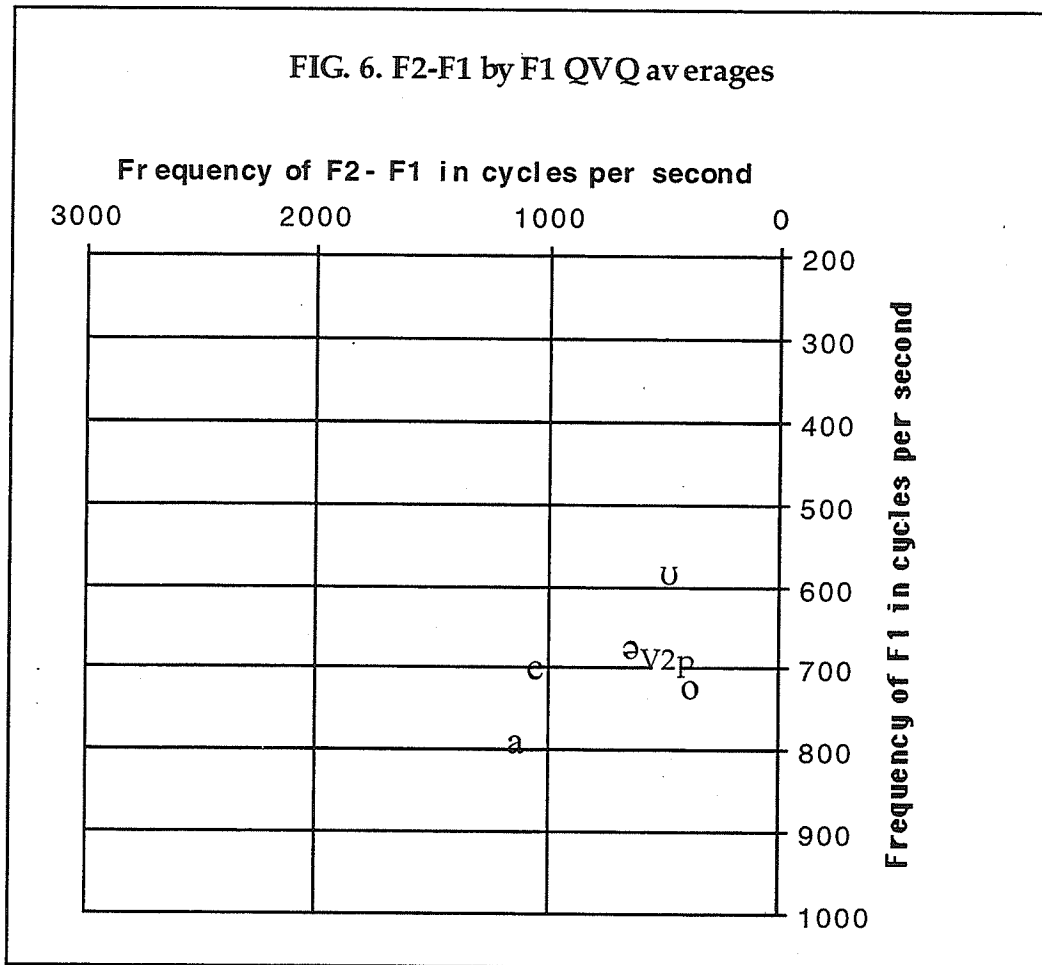


Figure 6 shows that, when both preceded and followed by uvulars, V2p is actually closer to /ə/ than to either /o/ or /u/.



Now we will consider the acoustic correlates of lip rounding since the /ə/ and /u/ in uvular environments are described (Jerue, et al., 1993) as having the same quality except for lip rounding. Rounding is diachronically important to Athabaskan languages. Proto Athabaskan is reconstructed with consonantal rounding distinctions for velars (e.g. *k,*k^w), and uvulars (e.g. *q,*q^w) (Krauss 1964:120). Many instances of Deg Xinag /u/ can be traced to */ə/ or */a/ adjacent to a uvular.

The universal acoustic correlates of rounding are not as clear as are those for height and backness. Lindblom and Sundberg (1971) claim that lip rounding lowers all formants, especially F3 in front vowels, and F2 in back vowels. Crothers (1978) points out that lip rounding with back vowels or spreading with front vowels reinforces the acoustic correlates of tongue position.

resonating in the space between a front tongue constriction and the lips is further heightened by spreading the lips thus shortening this resonating chamber. A lower F2 resonating in the space between a back tongue tongue constriction and the lips is further lowered by rounding the lips thus lengthening this resonating chamber.⁶ Schwartz, et al, (1993) found that in [i] vs. [y] contrasts, F2 remained stable but the resonance associated with F3 could fall below that of F2. When this happened, both frequencies appeared to become lower with rounding when, in fact, only F3 was falling. These studies indicate that we should expect lower formant values (but not clearly which formants) for a rounded vowel versus a similar vowel that is not rounded.

Figure 7 gives F1, F2, and F3 averages for each of the vowels in coronal and uvular environments. Here we find all the F1 and F2 values lower for /u/ than /ə/ (in corresponding environments, i.e., CuC vs. CəC) as we would expect if /u/ is the rounded equivalent of /ə/. The F3 comparisons, however, are not so robust; only the CVQ F3 is lower for /u/ than /ə/.

Comparing the data in Figure 7 for a clear correlation of formant lowering and rounding associated with uvularization does not produce robust results. The F1 of /e/, /ə/, /o/, /a/, and V2p is higher in post uvular than post coronal environments. The F2 of /ə/, /u/, and V2p is lower in post uvular than post coronal environments but it is slightly higher for /o/. F3 for /a/, /o/, /u/, and 2p, contrary to prediction, is **higher** in post uvular compared to post coronal environments. Looking at the vowels in pre uvular settings yields an equally murky outcome, though F2 and F3 are generally lower for vowels in pre-uvular compared to pre-coronal environments.

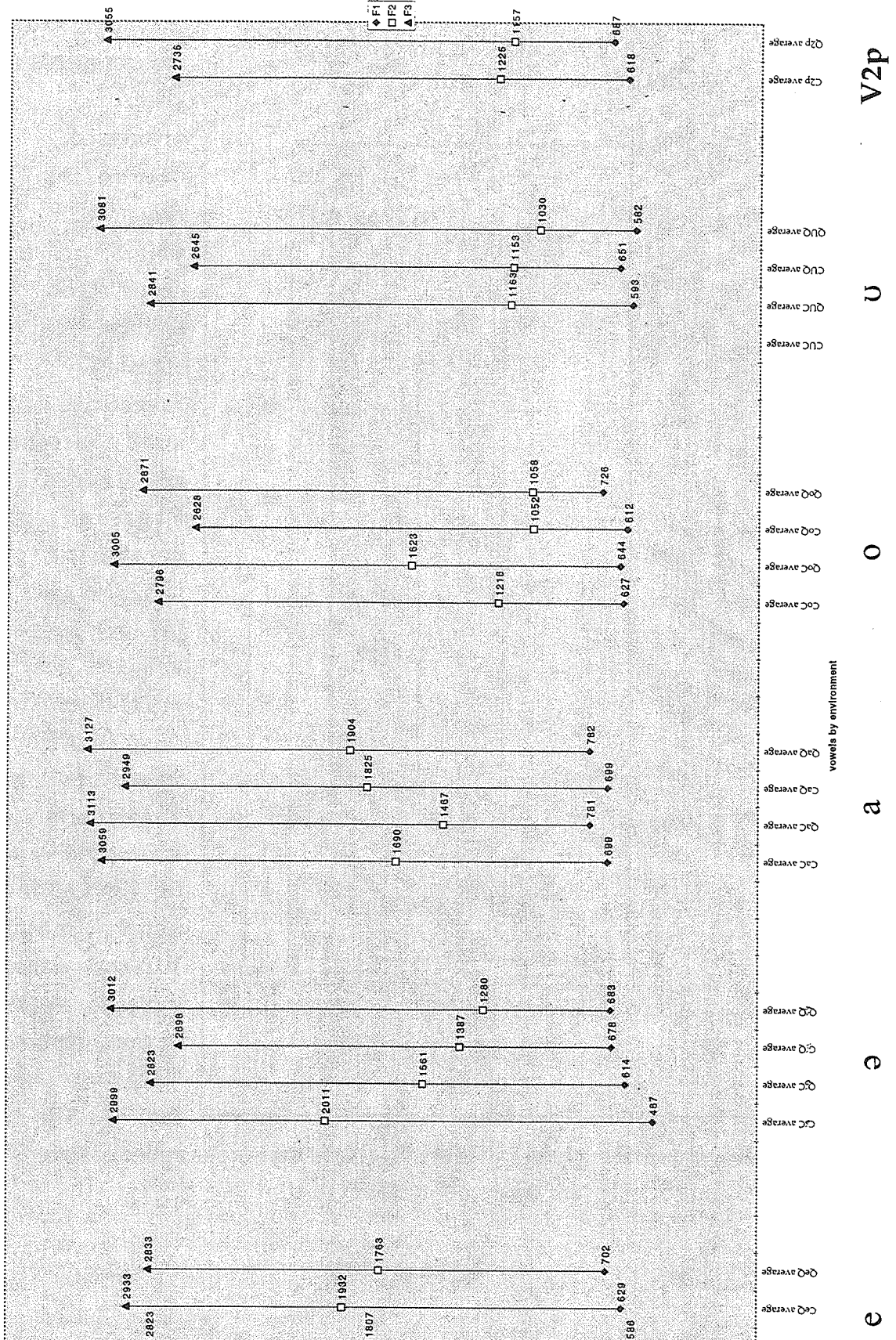
Measurements of F1, F2, and F3 do not clearly show that any of the same-vowel pairs, on the basis of uniformly lower formants, has one variant more rounded than the other. Nor does it clearly show that /u/ is more rounded than /ə/ as described in Jerue et al. (1993).

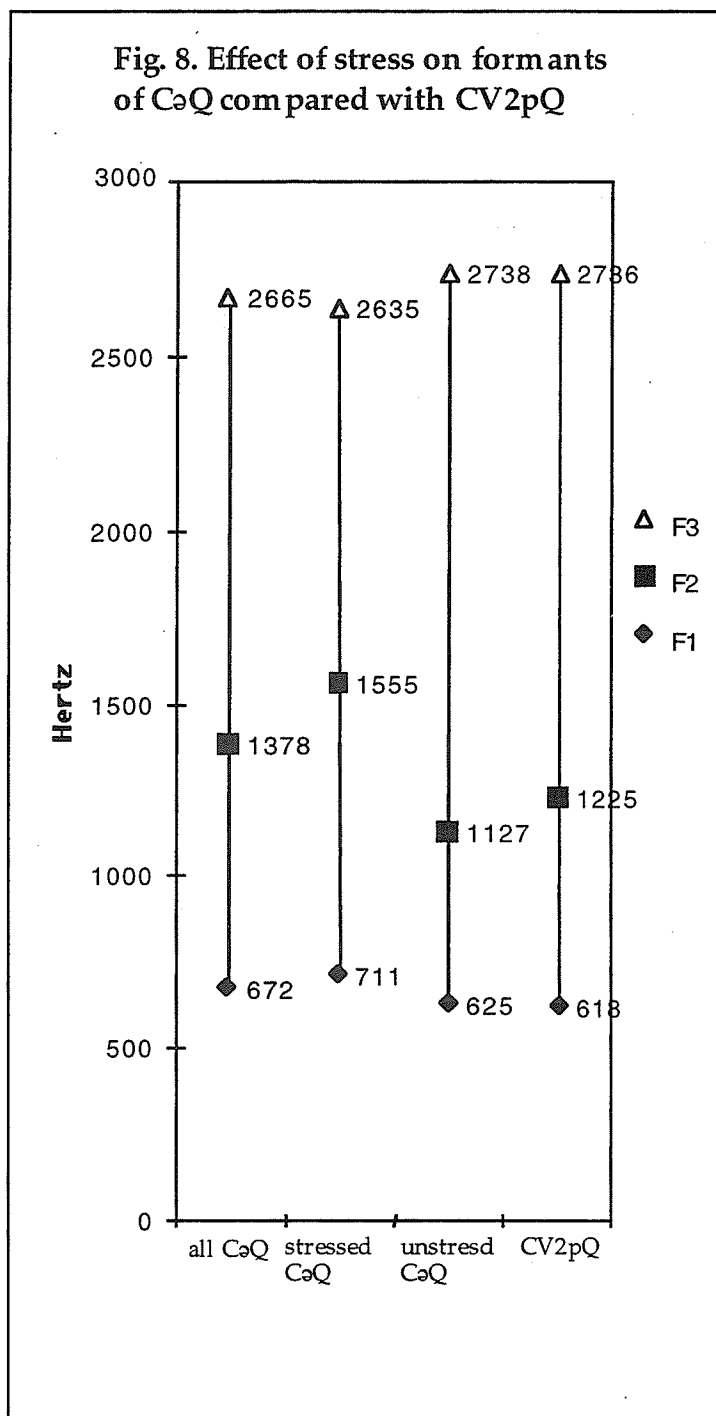
What the graphic display of formant averages in Figure 7 **does** suggest is that the overall profile of V2p is more like that of /u/ than either /o/ or /ə/. (Compare /u/, /o/, and /ə/ before Qs to V2p.)

To determine whether stress affects formant structure, further subdivision of

⁶ This is why back vowels are commonly rounded but front vowels only infrequently so. From the UCLA Phonological Segement Inventory Database, Maddieson (1984) lists 44 back-rounded vowels found in hundreds of languages but only 16 front-rounded vowels in approximately 40

Fig. 7. F1, F2, and F3 averages for each vowel in each environment





the vowel environments into stressed and unstressed categories is required. This yields, in most cases, too few tokens to be statistically revealing. One such attempt, illustrated in Figure 8, comparing the CəQ data to CV2pQ suggests, however, that not only does stress seem to affect the values of all three formants, but the unstressed version of CəQ more nearly matches V2p than does the stressed version. In fact, unstressed CəQ more nearly matches V2p than

any of the vowel profiles given in Figure 7. (Note that the number of tokens for unstressed CəQ is 5, only half the target number of 10.)

3.2 DURATION

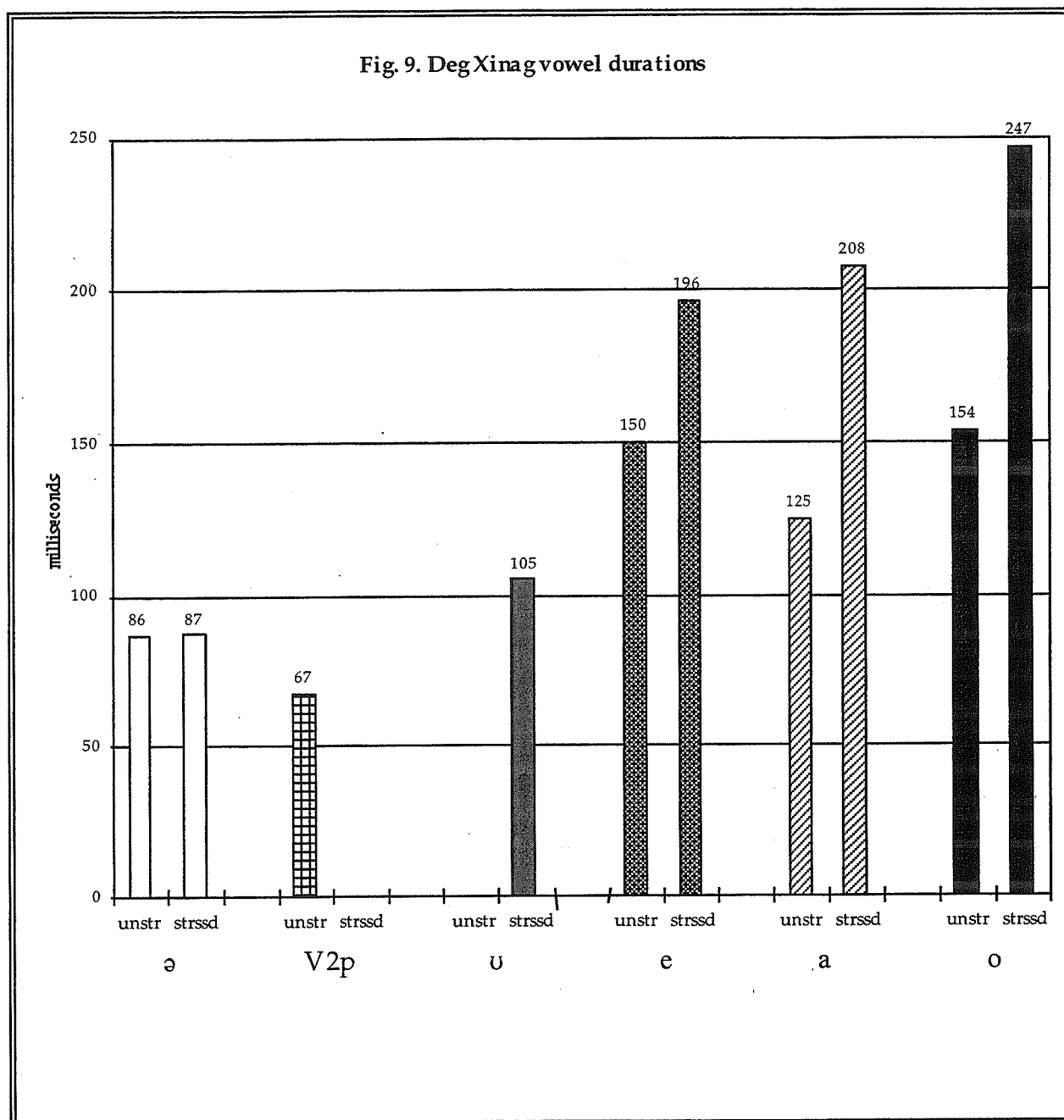
is described as long and the /u/ as short. If duration measurements of these two vowels prove to be significantly different from each other, then the duration of the questionable vowel, V2p, can be matched against them to identify it. Duration measurement averages for all the Deg Xinag vowels are reported here.

Current studies suggest that, in addition to formant interval relationships, vowel recognition is greatly improved by vowel duration information, even for languages in which vowel length is not contrastive. In American English vowel recognition experiments, Strange (1989a, b) compared results of unmodified versus modified consonant-vowel-consonant stimuli. She found that when listeners were presented with unmodified syllables, vowels were identified at a 6% error rate. When steady state formant information was attenuated to silence (silent steady state) but onglides and offglides were retained and duration information remained intact (i. e., consonants were retained, the vowel was silent but the silence was of the original vowel length), there was a vowel identification error rate of 7%. Similarly, when only the steady state formants were given without onglides and offglides but again, duration information remained intact, there was a 6% error rate. . On the other hand, when vowel length was neutralized, error rates rose to 14% for the silent steady state and 33% for the steady state without on and offglides. As a result, Strange proposes a 'dynamic specification model', one in which vowel duration is intrinsic to vowel recognition, as an improvement over solely formant ratio models.

Although Deg Xinag vowels are characterized as either short or long, length is not distinctive phonemically in the sense that there are no contrasting phonemes of the same quality that differ only in length. This is the case for Proto-Athabaskan. Figure 9 gives the average duration for each DX vowel, both stressed and unstressed. There are no examples in the corpus of unstressed /u/, and only one of stressed V2p, not enough to evaluate.⁷

⁷ The example of stressed V2p is found in the word *duxliquth* /dúχləqùθ/ 'you all are coughing', where the first syllable has the greatest amplitude of any syllable in the word. (Here ´ marks primary stress and ` marks secondary stress.) It is often the case that the stem (usually the last syllable) receives primary stress but here the amplitude of the first syllable is greater than that of the stem. The correlates of stress in Deg Xinag have not yet been determined so I remark only on

Fig. 9. DegXinag vowel durations



To determine the probability that the difference between duration averages for various vowel pairs was due to chance, a two-tailed t-Test was performed for each vowel pair. These results, given in Table 3, show that there is a significant difference between the stressed and unstressed pairs for the full vowels /a/ and /o/ but not for /e/.

For the reduced vowels, /u/ and /ə/, we face the problem that there are no unstressed /u/'s in the data against which to compare. For /ə/ there is no

significant difference between the stressed and unstressed duration averages, perhaps because this vowel is intrinsically so short that it cannot be significantly shortened in unstressed position. There is no significant difference in the durations of unstressed /ə/ and V2p so statistically, based on duration only, they could have the same identity.

Table 3. Probability that differences in pairs of vowel duration measurements are due to chance.	
Vowels	p value, two-tailed
stressed /o/ vs. unstressed /o/	.004***
stressed /a/ vs. unstressed /a/	.00000008***
stressed /e/ vs. unstressed /e/	.27
stressed /ə/ vs. unstressed /ə/	.67
stressed /u/ vs. unstressed V2p	.00003***
unstressed /ə/ vs. unstressed V2p	.08

Anything over .05 is non-significant. ** = highly significant difference ($p < .01$). *** = very highly significant difference ($p < .001$).

4. DISCUSSION

4.1 V2P IDENTITY

The various analyses above are summarized in Table 4 below. Here we see that, although there is some evidence that matches V2p with /o/, /ə/, and /u/, the strongest correlations for formant structure are among V2p and /u/ and /ə/. But certainly there are many cases where V2p has the same vowel quality as /o/ due to matching formant structure. So V2p has vowel quality which is sometimes the same as /o/, /u/, and /ə/.

The duration correlation narrows down the field of possible matches for V2p since /o/ is the longest of the Deg Xinag vowels and V2p is the shortest. Thus even though V2p may have the same quality as /o/, V2p is much too short to be /o/. The average for V2p is even shorter than the average for /ə/, the shortest of the vowels. There is no significant difference between V2p and /ə/. V2p may be /ə/, matching its quality and duration.

Or we can hypothesize that V2p is unstressed /u/ although we have no examples of unstressed /u/ against which to compare V2p. What we can do is compare V2p with stressed /u/ to see whether these two pattern like the stressed and unstressed versions of the other vowels. Like all the other same-vowel pairs, V2p is shorter than stressed /u/. As with the T-tests on /o/'s and /a/'s, the average duration of V2p is significantly different from that of stressed /u/. V2p may be /u/, matching its quality and duration.

The fact that V2p does match each of the three vowels /o/, /u/, and /ə/ along some of the parameters measured, accounts for the variation in its transcription.

Table 4. Summary of analyses matching V2p to /o/, /u/, and /ə/.

	F2-F1 by F1 averages		F1, F2, F3 averages		Duration
	CVQ Figure 5	QVQ Figure 6	CVQ Figure 7	QVQ Figure 7	Figure 9
/V2p/ = /ə/	close	closest	close	close	closest
/V2p/ = /u/	closest		closest	closest	close
/V2p/ = /o/	closest	close			

Another possibility is that V2p is no vowel at all. Since its average duration is 20 milliseconds shorter than the average of the shortest vowel, unstressed /ə/, the short vowelish space in the second person plural morpheme before /χ/ may be just that, a short vowelish space between consonants with no distinct vowel identity.⁸ Its features would be [+vowel, +round]. There is historical evidence that this may be the case. The Proto-Athabaskan form given as Figure 10, a. below is a labialized voiceless uvular fricative.

In some other Athabaskan languages the second person plural morpheme surfaces as listed in Figure 10, b.-o.

⁸ There may languages genetically unrelated to Deg Xinag with this type of very reduced duration vowel. Crothers (1978) cites Kuipers's data from Kabardian (Caucasian) as having two very short vowels, /i/ and /ə/ ((as well as five longer vowels). He describes the /i/ as, "...little more than a transitional sound between consonants,... regularly lost under certain conditions" (p 108). (Maddieson (1984) lists only one short vowel in Kabardian, /i/.) Coincidentally, Kabardian has, like Deg Xinag, a large phonemic consonant inventory. Kabardian has forty-eight consonants (Maddieson 1984) and Deg Xinag has fifty-two. Compare also Abkhaz, (Caucasian)

Figure 10. The phonemic shape of the Athabaskan morpheme 'second person plural'.

LANGUAGE	2p	SOURCE
a. Proto-Athabaskan	*χ ^w -	(Krauss and Leer 1981:197).
b. Koyukon	υχ	(Kari p.c.)
c. Dena'ina	əχ	(Kari p.c.)
d. Ahtna	ɔh	(Kari 1990:57)
e. Slave	ah-	(Rice 1989:481)
f. Central Carrier	υχ	(Kari p.c.)
g. Chipewayan	uh-	(Li 1946:411)
h. Galice	o?o	(Hoijer 1966:325)
i. Tolowa	u?	(Bright 1964) also o? (Siri Tuttle p.c.)
j. Mattole	oh	(Li 1930:69)
k. Tututni	o?-	(Golla 1976)
l. Navajo	oh-	(Kari 1976:21)
m. Tanana	ah	(Tuttle p.c.)
n. Eyak	əχ	(Krauss and Leer 1981:45)
o. Babine/Witsu Wit'en	χ ^w	(Hargus 1991)

All the Proto-Athabaskan reflexes in Figure 10, b.-n. have in common a back vowel and a back consonant.

The Deg Xinag 2p representation seems to surface, like Babine/Witsu Wit'en, Figure 10, o., as the exact reflex of the reconstructed Proto Athabaskan second person plural form. Rather than /oχ/ or /υχ/, the underlying morpheme might be /χ^w/, with labialization surfacing before the consonant. Along the same lines, Kari (1990: 652) suggests that the Ahtna areal prefix /ko/ may actually be vowelless /k^w/ with the velar labialization surfacing as "rounded o" after the consonant. I suggest a similar case for the Deg Xinag second person plural morpheme. Recall that ff. 2 mentions transcription variety in the vowel in the Deg Xinag areal prefix. Analysis of the measurements of the areal prefix vowel could support the 'vowelish' vowel hypothesis above if the areal prefix vowel also turns out to be very short and of neutralized vowel quality. This implies that as the vowelless morphemes with labialized consonants of Proto-Athabaskan develop into the reflex languages, these morphemes may not develop full vowels even if labialized consonants are no longer contrastive in the

How should writers of Deg Xinag spell the second person plural morpheme? Since a labialized uvular series is not part of the current orthography,⁹ I will select one of the vowels for V2p. On the basis of vowel quality we have seen that /ʊ/, orthographically 'u', /ə/, orthographically 'i', and /o/, orthographically 'o', are fairly neutralized in uvular environments. See the appendix for the amount of overlap in their vowel space areas. On the basis of duration though, V2p is closer to what we might expect of an unstressed /ʊ/ than an unstressed /o/. Thus, I suggest that the best orthographic representation is 'ux'. 'u' is the orthographic symbol for /ʊ/.

4.2 VOWEL SPACE

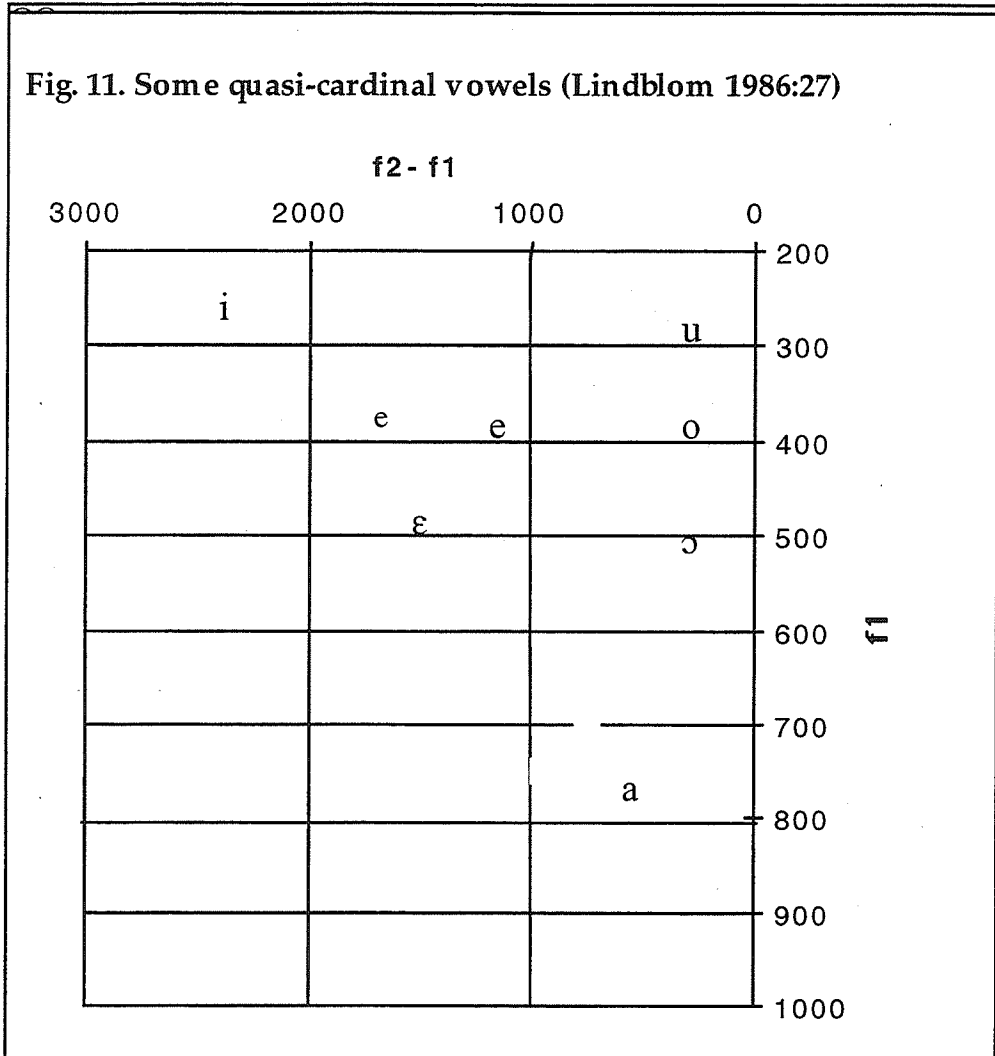
Although strong coarticulation effects resulting from the phonetic environments measured here are evident for each of the vowels,¹⁰ this investigation appears to bear out earlier descriptions of the Deg Xinag vowel system (Kari 1978, Leer 1979, Krauss and Golla 1981, Hargus and Taff 1991) as an unusual one in that it lacks contrastive use of high vowel space. Virtually all other vowel inventories use high vowel space contrastively.

Compare the Deg Xinag use of high vowel space with some of the quasi-cardinal vowel measurements reported in Lindblom 1986 given here as Figure 11. Looking at F1 measurements for the high vowels, we see that Lindblom's highest front vowel, /i/, is 230 Hertz away from the highest Deg Xinag vowel, /ə/ (using the average measurement for /CəC/ which is shown in Figure 3). The Lindblom /i/ is 100 Hertz away from the highest data point for all /ə/s measured in this data (see Figure 14). Looking at the back high vowels, we see that Lindblom's highest back vowel, /u/, is 300 Hertz away from the highest Deg Xinag back vowel, /ʊ/ (using the average measurement for /QʊC/ which is shown in Figure 4). Lindblom's /u/ is 200 Hertz away from the highest data point for all /ʊ/s measured in this data (see Figure 17). I give these comparisons to show that what we think of as /i/ and /u/, here represented by cardinal vowel measurements, are not reported for Deg Xinag **phonemically** as averages of the measured data, nor **phonetically**, taking each data point into

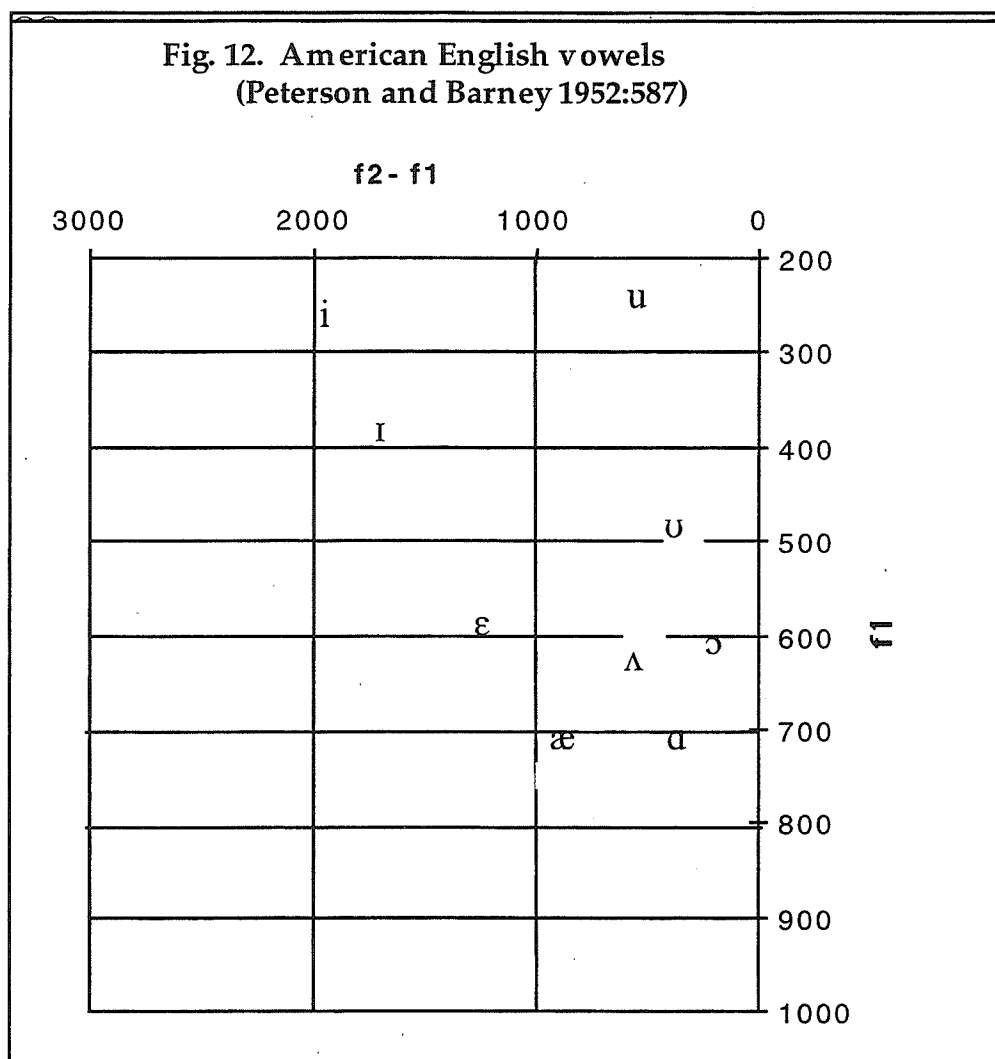
⁹ The Deg Xinag alphabet is; ', ch, ch', d, ddh, dh, dl, dr, dz, g, gg, gh, h, j, k, k', l, l', m, m', n, nh, n', ng, ngh, ng', p, q, q', s, sh, sr, t, t', th, tl, tl', tr, tr', ts, ts', tth, tth', v, x, y, yh, y', z, zr, a, e, i, o, u.

¹⁰ See the appendix for formant frequency plots of each vowel by phonetic environment. These

consideration.



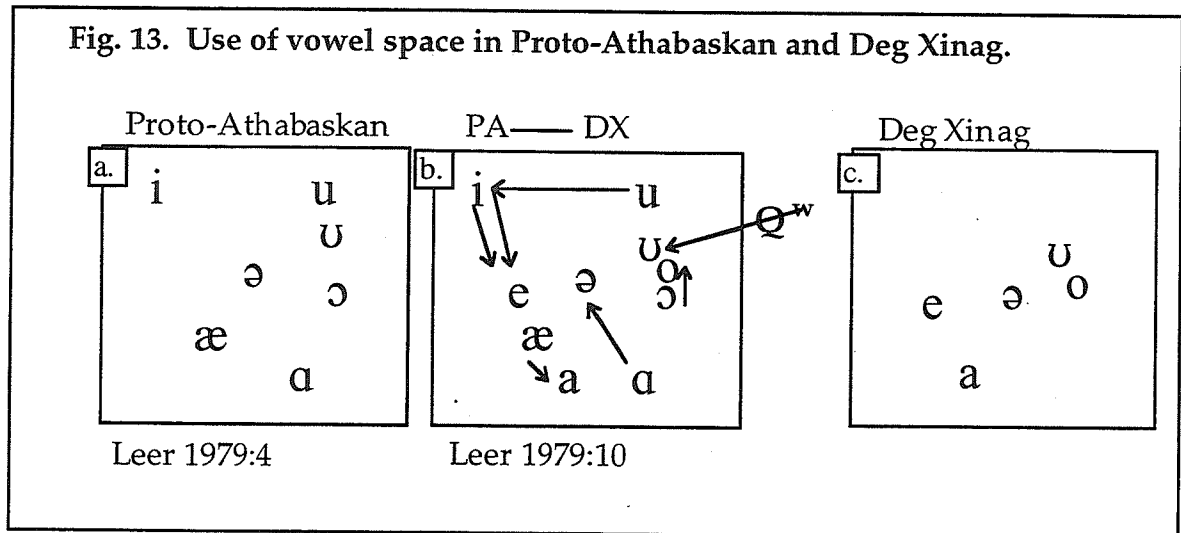
We get the same result by comparing Deg Xinag vowels to American English (mid-Atlantic) vowels as described in Peterson and Barney (1952) given here as Figure 12, since the high vowel F1 measurements are almost the same as Lindblom's.



Vowel space theory (e.g. Liljencrants and Lindblom 1972, Crothers 1978, Disner 1984, Lindblom 1986) predicts that, regardless of the number of vowels in a particular language's phonology, these vowels arrange themselves to fill all the acoustic space available for vowel perception. The research reported here shows, on the contrary, that the vowel system of Deg Xinag, like its cousin language Navajo, does not 'fill up' the available vowel space (McDonough and Ladefoged

1994).¹¹ Deg Xinag vowels as described here in acoustic terms of their F1 and F2 do not make use of either the highest front or back vowel space. Although /ə/ appears more high and front in a CəC environment than it does in other environments, Deg Xinag does not have any phonemic high vowels. Crothers (1978) does not report any vowel system in which all the vowels are closely bunched in a small area as are the DX vowels reported on here. However, Ladefoged and Maddieson (1996: 286) note that the Chadic language Margi has only /i/ and /a/; the Australian language Arrernte has only /ə/ and /a/; the Caucasian languages Ubykh and Abkhaz have only /ə/ and /a/; and the Caucasian language Kabardian has only /i/, /ə/, and /a/ in their vowel inventories. So we see that not all languages disperse their vowel phonemes to 'fill up' vowel space. Ladefoged and Maddieson do indicate though that at least for Arrernte and Kabardian, the vowels do disperse **phonetically** to the edges of vowel space. This does not seem to be the case in Deg Xinag.

If we compare Proto-Athabaskan to Deg Xinag vowels we see that the latter have moved towards the center of vowel space. Figure 13., a. illustrates the Proto-Athabaskan vowels. The arrows in Fig. 13, b. indicate the direction of vowel shift of each vowel from Proto-Athabaskan to Deg Xinag. Figure 13, c. gives the Deg Xinag vowels.



5. CONCLUSION

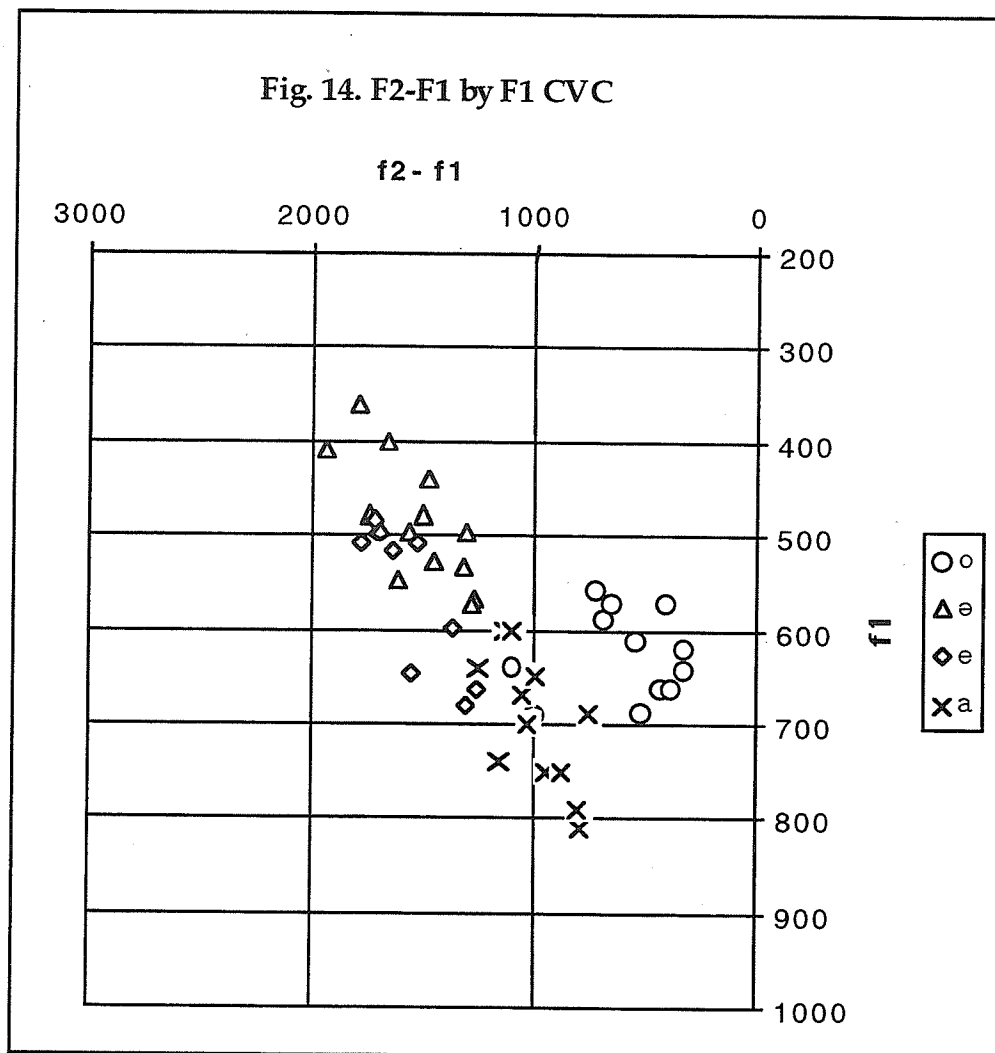
What is the answer to the question that triggered this study: Can acoustic measurements of the Deg Xinag vowel in the second person plural morpheme contribute to its phonemic disambiguation? We can answer that measurements of the first three formants as well as vowel duration have accounted for the controversy over the identity of V2p since it has characteristics of several of the other vowels. In a forced choice answer among /o/, /ə/, and /u/, I have selected /u/ as the most likely candidate for V2p since they are the most similar. But, on the basis of its very short duration, I have questioned the 'full vowel' status of V2p, suggesting that as the reflex of the labial component of a historical labialized uvular it did not develop into a full vowel. Support for this analysis could be found in measurements of the vowels of other morphemes that were historically vowelless labialized consonants, e. g., the areal prefix.

The measurements made for this study have corroborated previous descriptions by linguists (Kari 1987, Leer 1979, Kraus and Golla 1981, Hargus and Taff 1991) who note no high vowels in the phoneme inventory of Deg Xinag.

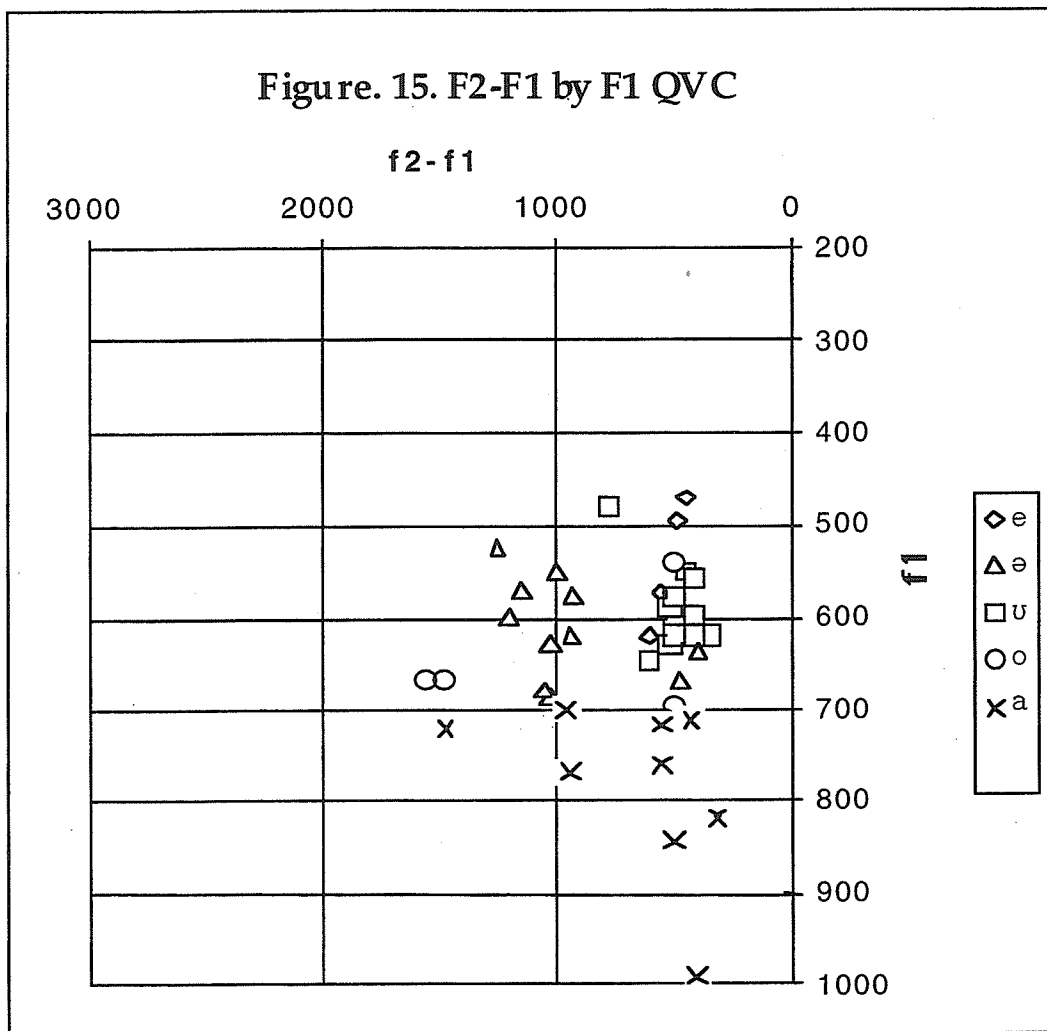
APPENDIX

Each scatterplot in Figures 14 through 17 represents one of the four measured vowel environments. The distribution of frequency measurements for all the data analysed is plotted showing F_2-F_1 against F_1 .

The chart of vowel frequencies in the CVC environment given as Figure 14 does not contain data for /u/ or V2p since by definition, /u/ occurs almost exclusively pre and post-uvularly and the V2p always precedes /χ/. The data has an unsurprising distribution, /e/ mid front, /o/ mid-back, /a/ low, and /ə/ mid central and high front under influence from the coronal consonants. Each of the vowels has its own space and there is not much overlap among vowels.



In the QVC environment charted in Figure 15 the /e/ appears far enough back to overlap with /o/ and /u/. /ə/ is central.



In the Coronal V Uvular environment, illustrated in Figure 16, the distinction between vowels becomes less clear. The area for /o/ is almost entirely within the area for /ə/, which itself is almost entirely within the area for /u/.

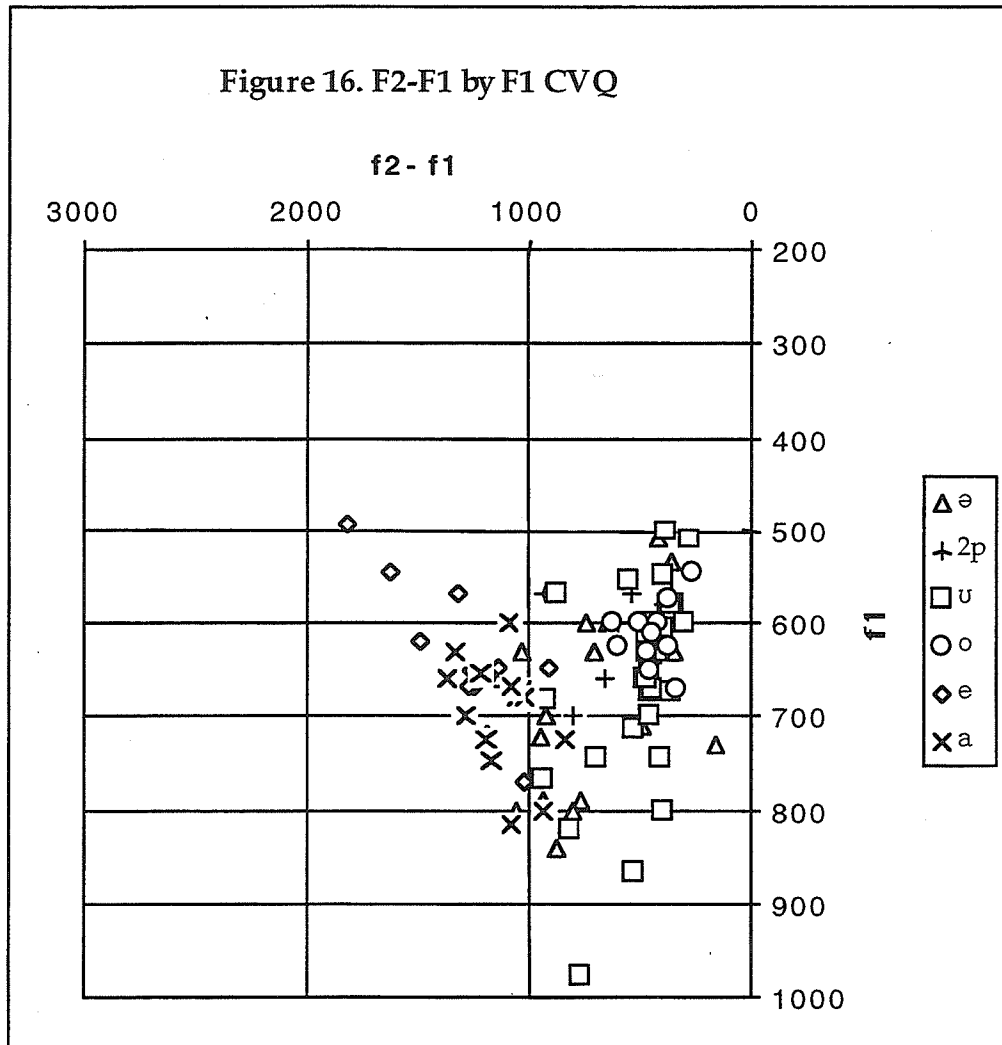
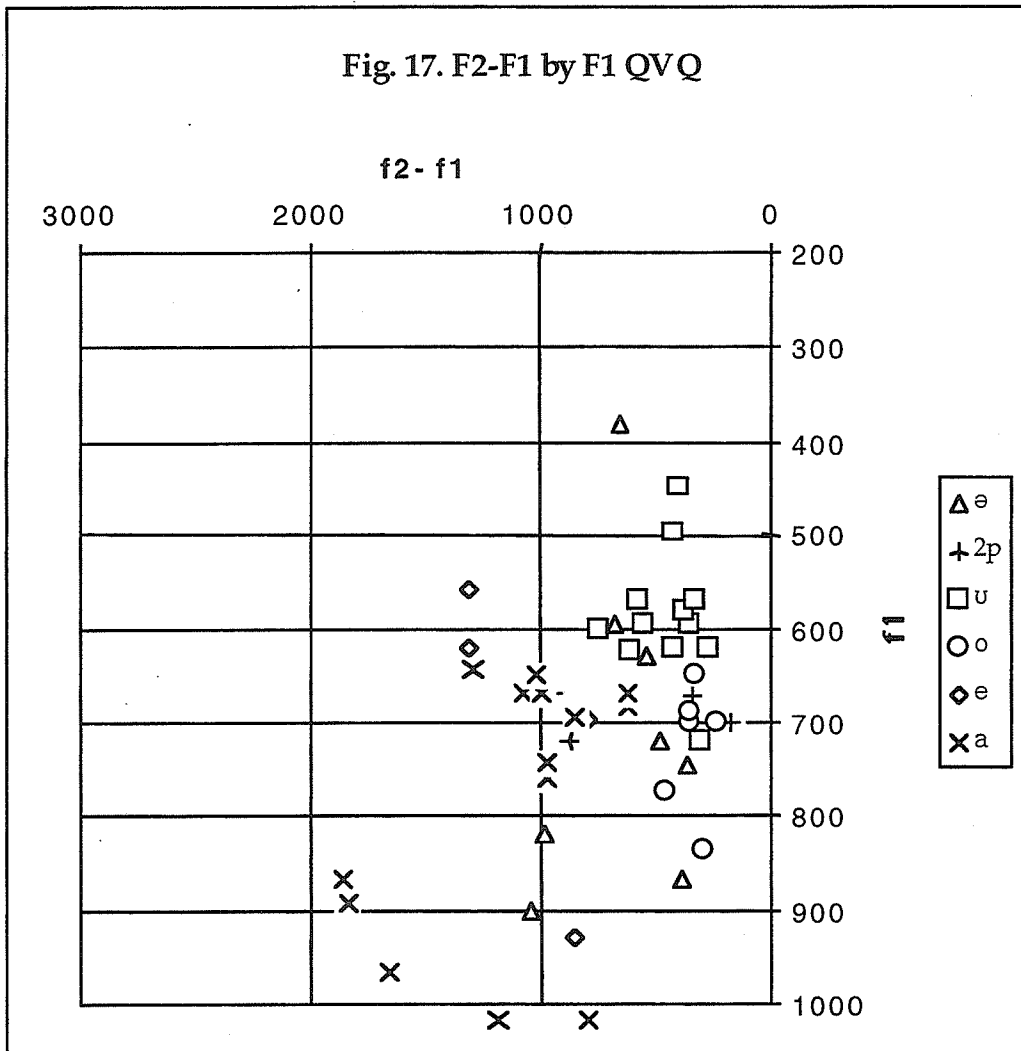


Figure 17 illustrates the neutralization possibilities for vowels in a QVQ environment.



REFERENCES

- Antoine, Francesca, et al. 1974. *Central Carrier Bilingual Dictionary*. Fort St. James, British Columbia: Carrier Linguistic Committee.
- Bright, Jane. 1964. Phonology of Smith River Athapaskan (Tolowa). *International Journal of American Linguistics* 30: 101-107.
- Crothers, John. 1978. Typology and universals of vowel systems. *Universals of Human Language 2: Phonology*, ed. by Joseph H. Greenberg, 93-152. Stanford: Stanford University Press.
- Golla, Victor. 1970. *Hupa Grammar*. University of California, Berkley. Ph.D. thesis.
- _____. 1976. Tututni (Oregon Athapaskan). *International Journal of American Linguistics* 42: 217-227.
- Hargus, Sharon. 1991. The disjunct boundary in Babine/Witsu Wit'en. *International Journal of American Linguistics* 57: 487-513.
- Hargus, Sharon, James Kari and Alice Taff, compilers. (in preparation) *Deg Xinag verb paradigms*.
- Hargus, Sharon and Alice Taff. 1991. *Prefixal Dependencies in Deg Xinag*. University of Washington, MS.
- Hewitt, B. G. 1979. *Abkhaz. Lingua Descriptive Studies, 2*. Amsterdam: North-Holland Publishing Company.
- Hoiijer, Harry. 1966. Galice Athapaskan: a grammatical sketch. *International Journal of American Linguistics* 32: 320-27.
- Jerue, Alta, Hannah Maillelle, Alice Taff, and Sharon Hargus. 1993. *Deg Xinag Dindlidik*. Fairbanks: Alaska Native Language Center.
- Kari, James. 1976. *Navajo Verb Prefix Phonology*. New York: Garland Publishing, Inc.
- _____. 1978. *Deg Xinag: Ingalik noun dictionary*. Fairbanks: Alaska Native Language Center.
- _____. 1990. *Ahtna Athabaskan Dictionary*. Fairbanks: Alaska Native Language Center.
- Koopmans-van Beinum, F.J. 1973. Comparative phonetic vowel analysis. *Journal of Phonetics* 1: 249-61.
- Krauss, Michael E. 1964. Proto-Athabaskan-Eyak and the problem of NaDene: the phonology. *International Journal of American Linguistics* 30: 118-31.

- Krauss, Michael E. and Victor K. Golla. 1981. Northern Athapaskan languages. Handbook of North American Indians 6: 67-85. Washington, D.C.: Smithsonian Institute.
- Krauss, Michael E. and Jeff Leer. 1981. Athabaskan, Eyak, and Tlingit sonorants. Fairbanks: Alaska Native Language Center.
- Latimer, Richard Murray. 1978. A study of Chilcotin phonology. Master's thesis, University of Calgary.
- Leer, Jeff. 1979. Proto-Athabaskan verb stem variation, part one: phonology. Fairbanks: Alaska Native Language Center.
- Li, Fang-Kuei. 1930. *Mattole: An Athabaskan Language*. Chicago: University of Chicago Press.
- _____. 1946. Chipewyan. in *Linguistic Structures of Native America* ed. by Cornelius Osgood. New York: Viking Fund Publications in Anthropology 6: 398-423.
- Liljencrants, Johan and Björn Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48, 4: 839-62.
- Lindblom, Björn and J Sundberg. 1971. Acoustical consequences of lip, tongue, jaw and larynx movement. *Journal of the Acoustical Society of America* 50: 1166-79.
- Lindblom, Björn. 1986. Phonetic universals in vowel systems. in *Experimental Phonology*, ed. by J.J. Ohala and J.J. Jaeger, 13-34. New York: Academic Press.
- Kuipers, A. H. 1960. *Phoneme and Morpheme in Kabardian*. The Hague: Mouton.
- Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.
- McDonough, Joyce, Peter Ladefoged, and Helen George. 1993. Navajo vowels and phonetic universal tendencies. *Fieldwork studies of targeted languages, UCLA Working Papers in Phonetics* 84: 143-50.
- McDonough, Joyce and Peter Ladefoged. 1993. Navajo stops. *Fieldwork studies of targeted languages, UCLA Working Papers in Phonetics* 84: 151-64.
- Peterson, Gordon E. and Harold L. Barney. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24, 2: 175-84.
- Rice, Keren. 1989. *A Grammar of Slave*. Berlin: Mouton de Gruyter.

- Schwartz, Jean-Luc, Denis Beautemps, Christian Abry and Pierre Escudier. 1993. Inter-individual and cross-linguistic strategies for the production of the [i] vs. [y] contrast. *Journal of Phonetics* 21: 411-25.
- _____. 1989b. Evolving theories of vowel perception. *Journal of the Acoustic Society of America* 85, 5: 2081-87.
- Strange, Winifred. 1989a. Dynamic specification of coarticulated vowels spoken in sentence context. *Journal of the Acoustic Society of America* 85, 5: 2135-53.
- Tenenbaum, Joan. 1979. Morphology and semantics of the Tanaina verb. Columbia University. Ph. D. thesis.
- Thompson, Chad, Melissa Axelrod, and Eliza Jones. 1983. Han Zadlitłee: Yukon Koyukuk School District Koyukon Language Curriculum. Nenana, Alaska: Yukon Koyukuk School District.

Licensing and Scope Interpretation of *Wh*-phrases in Korean

Tai-Soo Kim

1. Introduction

In this paper, I will re-examine the literature dealing with question formation, especially the licensing of *wh*-words in Korean, Japanese, and English, and propose a feature system for Complementizers and functional morphemes involved in *wh*-question formation and scope interpretation of *wh*-words in the framework of the MP.

It has been a widely accepted generalization (I will call this '*wh*-Constraint.')

about the licensing of *wh*-words in Korean and Japanese that *wh*-words must be in some tight relationship with an interrogative morpheme (Kuno (1973), Saito (1985, 1992), Nishigauchi (1990), Yoon (1990), Kim (1991), Jung (1991), Watanabe (1991)).¹ It has been also argued that this *wh*-Constraint works as the *wh*-island which does not allow embedded *wh*-words to take the matrix scope. I will show that the *wh*-Constraint cannot fully apply to Korean *wh*-words, unlike Japanese. Korean has a well-developed system of sentential endings and they are tightly connected to a certain group of verbs. It seems that this tightness overwhelms the *wh*-Constraint

¹ The formal statements of the Constraint are as follows;

(i) *Wh*-word Constraint

a. An interrogative polarity item (IPI) must be governed by an interrogative marker (Q-morpheme) at LF. (Kim 1991: 219)

b. WH must be governed by [+Q]. (Nishigauchi 1990: 21)

Application of these constraints will be discussed in Section 2.

which might apply to Korean in the same fashion as Japanese, which is another language in which *wh*-movement takes place at LF. I will classify Korean verbs into four types and suggest some features for the sentential endings, and show that the *wh*-Constraint applies only to the embedded *wh*-words introduced by *think*-type matrix verbs. This will be a subcategorization of Complementizers. I will argue that the *wh*-Constraint concerns the SPEC-Head agreement in CP in the feature-checking theory and that only in this case the embedded *wh*-words show *wh*-island effect, but in many cases, Korean *wh*-words need not to be in the SPEC-Head agreement with Comps and do not show *wh*-island effect.²

In §2, I present data to which the *wh*-Constraint does not apply. In §3, English question formation will be briefly examined in terms of feature setting, and properties of Korean Complementizers will be explored for question formation and *wh*-word licensing. §4, I will classify four types of verbs which show different behaviors in the process of *wh*-word licensing and scope interpretation of *wh*-words. In the final section a conclusion will be presented including residual problems.

² Rizzi (1990) argues that even in English *wh*-words are in the SPEC-Head agreement with null-Comp head. If we want to make this argument completely compatible with the Constraint of *wh*-words, we have to assume that English null-Comp head has [+WH_v] feature (to be explained in Section 3.2.) and we also have to assume a null-interrogative morpheme which has [+WH_v] feature in the case of Korean stand-alone *wh*-words. But I am not sure whether Korean has further motivation to establish the null-interrogative morpheme, and I just set aside this possibility.

2. Licensing of *Wh*-words in Korean and Japanese

2.1. *Wh*-word Constraints

In the literature, it has been argued that Korean and Japanese *wh*-words need a certain special morpheme to be licensed in sentences. Let us consider the Korean data in (1a). Unlike English, in these indirect *wh*-questions, a special Q-morpheme *-nunka* is used. As shown in (1b), when the *wh*-word *encey* ‘when’ is used with the sentential ending *-ta-ko*, the sentences are out. Japanese shows exactly the same phenomenon, as illustrated in (3). In these Japanese indirect questions, the Q-morpheme *-ka* is used. These Q-morphemes, Korean *-nunka* and Japanese *-ka*, are used for direct questions as well, as shown in (2) and (4).³

- (1) a. na-nun Jihi-ka **encey** ka-ess-**nunka** sayngkakha / al / kiekha -n-ta.
I-Top Jihi-N **when** go-Pa-**Int** think / know / remember-Pr-Dec
‘I think / know / remember when Jihi went away.’
- b. *na-nun Jihi-ka **encey** ka-ess-**ta-ko** sayngkakha / al / kiekha -n-ta.
I-Top Jihi-N **when** go-Pa-**Dec-Sub** think / know / remember-Pr-Dec
‘I think / know / remember when Jihi went away.’
- (2) a. Jihi-ka **encey** ka-ess-**nunka**?
Jihi-N **when** go-Pa-**Int**
‘When did Jihi go away?’
- b. Jihi-ka **cipey** ka-ess-**nunka**?
Jihi-N home go-Pa-**Int**
‘Did Jihi go home?’

³ Abbreviations used in this paper are as follows;

N: Nominative Marker	Dec: Declarative ending	Pr: Present Tense
A: Accusative Marker	Int: Interrogative ending	Pa: Past Tense
H: Honoric Marker	Imp: Imperative ending	Op: Operator
T: Topicalization Marker	Coh: Cohortative ending	Su: Subordinator
P: Plural Marker		

For the sake of space largely irrelevant morphemes are sometimes just glossed with capital letters.

- (3) a. Boku-wa dare-ga kuru ka sir-i-mas-en / oboe-te-i-masu.
 I Top who-N come Int know not / remember-is
 'I {do not know / remember} who is coming.'
- b. *Boku-wa dare-ga kuru to sir-i-mas-en / oboe-te-i-masu.
 I Top who-N come Sub know not / remember-is
 'I {do not know / remember} who is coming.'
- (4) a. Dare-ga kuru ka?
 who-N come Int
 'Who is coming?'
- b. Saito-ga kuru ka?
 Saito-N come Int
 'Is Saito coming?'

From this observation, many linguists (Kuno 1973; Saito 1985 & 1992; Nishigauchi 1990; Yoon 1990; Kim 1991; Watanabe 1991; Jung 1992; among others) conclude that in Korean and Japanese *wh*-words must be in some tight syntactic relationship with the interrogative morpheme *-nunka* (*-ka* in Japanese). Specifically, Kim proposes a language-specific constraint as stated in (5a), and similarly, Nishigauchi proposes (5b).

(5) *Wh*-word Constraint

- a. An interrogative polarity item (IPI) must be governed by an interrogative marker (Q-morpheme) at LF. (Kim 1991: 219)
- b. WH must be governed by [+Q]. (Nishigauchi 1990: 21)

Nishigauchi argues that this *wh*-Constraint exerts *wh*-island effect. That is, in (6a) the reason why the *wh*-word *dare* 'who' or *nani* 'what' cannot take wide-scope is due to the embedded Q-morpheme *ka*, while in (6b) *dare* takes only wide-scope since only the matrix Q-morpheme is available.

- (6) a. Tanaka-kun-wa dare-ga nani-o tabe-ta-ka oboe-te-i-masu-ka?
 Tanaka Top who-N what-A eat-Pa-Int remember-is Int
 'Does Tanaka know who ate what?'
 not 'For which x, x a person, does Tanaka know what x ate?'
 not 'For which y, y a thing, does Tanaka know who ate y?'
 (Nishigauchi 1990: 28)
- b. Sinbun-wa dare-ga erab-are-ta to tutae-te-i-masu-ka?
 newspaper-Top who-N elected-was Int report-is-Int
 'Who do the newspaper report was elected?'
 not 'Does the newspaper report who was elected?'
 (Nishigauchi 1990: 19)

Jung (1992) claims Korean shows a similar phenomenon. That is, in (7a) the *wh*-word *encey* takes only narrow-scope since the embedded clause has the Q-morpheme *-nya*, while in (7b) it takes only wide-scope since the Q-morpheme *nya* is in the matrix sentence.

- (7) a. Chelswu-nun Yenghi-ka encey ttena-ess-nya-ko mwul-ess-nya?
 Chelswu-Top Yenghi-N when leave-Pa-Int-Su ask-Pa-Int
 'Did Chelswu ask when Yenghi left?'
 not 'When_i did Chelswu ask Yenghi left t_i?' (Jung 1992: 53)
- b. Chelswu-nun Yenghi-ka encey ttena-ess-ta-ko malhay-ess-nya?
 Chelswu-Top Yenghi-N when leave-Pa-Dec-Su say-Pa-Int
 'When_i did Chelswu say Yenghi left t_i?
 not 'Did Chelswu say when Yenghi left? (Jung 1992: 53)

However, this *wh*-Constraint is not sufficient to explain the distribution and scope interpretation of Korean *wh*-words, as it will be shown in Section 2.1.

2.2. Problems for the *Wh*-word Constraint

Contrary to the expectation with the *wh*-Constraint, in Korean *report*-type verbs such as *pokoha*- ‘report’, *cenha*- ‘convey’, or *alli*- ‘inform’ allow *wh*-words in embedded clauses without an interrogative morpheme. That is, these verbs allow both types of embedded *wh*-clauses; constructions of *wh*-...-*ta-ko* which has no interrogative morpheme and those of *wh*-...-*nunci* which has the interrogative morpheme *-nunci*.

- (8) a. *Jihi-ka encey o-n-ta-ko pokoha / cenha / alli / -ess-ta.*
Jihi-N when come-Pr-Dec-Sub report / talk / inform -Pa-Dec
 ‘(I) reported/ convey/ inform when Jihi would come.’

- b. *Jihi-ka encey o-nunci pokoha / cenha / alli / -ess-ta.*

Int

More interestingly, cohortative or imperative verbs such as *ceyanha*- ‘suggest’ and *ceyuyha*- ‘propose’, or *myengha*- ‘order’ and *yokuha*- ‘request’ rather allow only embedded *wh*-clauses which are introduced with non-interrogative sentential endings; *ca* (cohortative ending), or *la* (imperative ending), as shown in (9) and (10), respectively.

- (9) a. *Minho-ka Jihi-eykey encey o-la-ko myengha-ess-ta.*
Minho-N Jihi-to when come-Imp-Sub order-Pa-Dec
 ‘Minho ordered Jihi when to come.’

- b. **Minho-ka Jihi-eykey encey o-nunci myengha-ess-ta.*

Int

- (10) a. *Minho-ka Jihi-eykey encey ka-ca-ko ceyanha-ess-ta.*
Minho-N Jihi-to when go-Coh-Sub suggest-Pa-Dec
 ‘Minho suggested Jihi when to go together.’

b.?* Minho-ka Jihi-eykey encey ka-nunci ceyanha-ess-ta.
Int

An important question arises at this point: Are the sentences in (8) through (10) real ‘indirect questions’ or ‘semi-questions’ in terms of Suñer (1991 & 1993)? If they are not indirect questions but pseudo or semi-questions, it will be a possible claim that the non-appearance of interrogative marker in them is a similar phenomenon to the fact that Spanish semi-questions do not allow *que*. According to Suñer, the complementizer *que* ‘that’ comes along with *wh*-words only in the real indirect questions in Spanish, but not in the semi-questions, as illustrated in (11).

- (11) a. Preguntaron (**que**) **quién** camina dormido. (Suñer’s (1993) (31a))
‘They asked (**th at**) **wh o** walks in his sleep.’
- b.*Explicaron **que** **cuál** cosecha se había dañado. (Suñer (1993) (50c))
‘They explained **th at** **wh ich** crop was damaged.’

In other words, to interpret the sentences in (8) in analogy with the Spanish sentences in (12), the sentences in (8b) which accompany the interrogative morpheme *-nunci* are indirect questions while the ones in (8a) which do not accompany the morpheme is semi-questions. The sentence in (12b) is a case of ‘overt lifting’ to the indirect question from the semi-question sentence in (12a). That is, it can be said that the Korean sentences in (8b) require the interrogative morpheme *-nunci* as (12b) requires *que*.

- (12) a. Repitieron **cu ándo** llegarían. (Suñer’s (1993) (38c))
‘They repeated **wh en** they would arrive.’

- b. Repitieron que a cuántos habíamos invitado. (Suñer's (1993) (38b))
 'They repeated (= asked repeatedly) that how many we had invited.'

However, this is hardly a possible analogy since it is hard to say that the Korean sentences in (8b) are indirect questions like the Spanish sentence in (12b). This is because the Korean matrix verbs in (8b) cannot be interpreted as question verbs even when they accompany the interrogative morpheme *-nunci* unlike the Spanish verb *repetir* 'repeat' which can be interpreted as 'ask repeatedly' when it comes along with *que*. This means that the Korean two types of sentences in (8a) and (8b) are all semi-questions. In the latter case, however, an interrogative morpheme is allowed in spite of its status of semi-question. And, I can distinguish no difference of meaning between the two types of sentences.⁴

As for scope interpretation of *wh*-words, I will clarify the interrogative morphemes used in question sentences in Korean. This is important since in the literature the interrogative morphemes *-nunci*, *-nya*, and *-nunka* are treated without distinction, and therefore, some confusion arises. For example, in my intuition the *wh*-word in (7a) can take a matrix-scope, and the *wh*-word in (7b) can take a narrow-scope, with appropriate intonation. Lee's (1982) intuition (from Nishigauchi 1990: 31-35) is the same as mine. He presents (13a) in which the embedded *wh*-words *nwukwu* 'who' or *mwues* 'what' can take wide scope. As a matter of fact, when the Q-morpheme *-nunka* is used instead *-nya-ko*, the *wh*-word takes only the narrow-

⁴ Another question is whether the sentences which do not accompany the interrogative morpheme are headless relative clauses. If they are headless relative clauses, the *wh*-words are not real question words and then we can say that it is the reason why they do not need an interrogative morpheme. I

scope, as shown in (13b). That is, the distinction of the two morphemes is important in the sense that one shows the *wh*-island effect but the other does not.

- (13) a. Bill-i nwukwu-ka mwues-lul ha-ess-nya-ko mwul-ess-nya?
 Bill-N who-N what-A do-Pa-Int-Su ask-Pa-Int
 ‘For which person *x*, Bill asked (you) for which thing *y*, *x* did *y*.’
 ‘For which thing *y*, Bill asked (you) for which person *x*, *x* did *y*.’
 ‘Bill asked, for which person *x* and for which thing *y*, *x* did *y*.’
 (Nishigauchi’s II-(47) taken from Lee (1982))
- b. Bill-i nwukwu-ka mwues-lul ha-ess-nunka mwul-ess-nya?
 Bill-N who-N what-A do-Pa-Int ask-Pa-Int
 ‘Bill asked, for which person *x* and for which thing *y*, *x* did *y*.’

There is another reason for this distinction. As shown in (14a) the Q-morpheme *-nya* licenses embedded *wh*-words only when the matrix verb is an *ask*-type verb. Whereas, the Q-morpheme *-nunka* allows any kind of verbs, as shown in (14b).

- (14) a. Jihi-ka encey ka-ess-nya ko mwul / *al / *kiekha -ess-ta.
 Jihi-N when go-Pa-Int Sub ask / know / remember-Pr-Dec
 ‘(I) asked / *knew / *remembered when Jihi went away.’
- b. Jihi-ka encey ka-ess-nunka mwul / al / kiekha -ess-ta.
 Jihi-N when go-Pa-Int ask / know / remember-Pr-Dec
 ‘I asked / knew / remembered when Jihi went away.’

Korean has another Q-morpheme which shows *wh*-island effect. In (15), the Q-morpheme *-nunci* shows the same *wh*-island effect and the *wh*-word takes only

am not sure about it. However, even when they are headless relative clauses, we need a feature system to license their appearance.

the narrow-scope. As for difference of the Q-morpheme *-nunci* from *-nya* or *-nunka*, *-nunci* can be used only in embedded clauses, as contrasted in (16).

- (15) ney-ka Bill-i nwukwu-lul po-ess-nunci kiekha-nya?
 you-N Bill-N who-A see-Pa-Int remember-Int
 'You remember, for which person *x*, Bill saw *x*.'
 Not 'For which person *x*, you remember Bill saw *x*.'
 (= Nishigauchi's II-41)
- (16) a. *Jihi-ka encey ka-ess-nunci?
 Jihi-N when go-Pa-Int
 'When did Jihi go away?' (intended reading)
- b. Jihi-ka encey ka-ess-nya/-nunka?
 Jihi-N when go-Pa-Int
 'When did Jihi go away?'
- c. Jihi-ka encey ka-ess-nunci mwul / sayngkakha / al / kiekha -ess-ta.
 Jihi-N when go-Pa-Int ask think / know / remember-Pr-Dec
 '(I) asked / thought / knew / remembered when Jihi went away.'

With these observations, I will claim that the three Q-morphemes *-nya*, *-nunka*, and *-nunci* must be treated differently, and I propose different Comp-features to license them in Section 4.

The distribution of *wh*-words can be summarized as in (17).

(17) Types of Verbs	wh-...nunci	wh-..ta-ko	wh-..nya-ko	wh-..la/ca-ko
	<i>-nunka</i>			
<i>think</i> -type	O	X	X	X
<i>report</i> -type	O	O	X	X
<i>ask</i> -type	O	X	O	X
Imp/Coh type	X/?	X	X	O
Q-Morphemes	Embedded Cl.	Matrix Cl.		
<i>-nunka</i>	O	O		
<i>-nunci</i>	O	X		
<i>-nya</i>	only with <i>ask</i> -type		O	

This distribution cannot be successfully captured with syntactic constraints such as those presented in (5). In the framework of MP in which the feature-checking system must reflect the properties of lexical items, if we explore the properties of Comp appropriately, all of these syntactic phenomena can be successfully explained as the result of the properties of the lexical items involved. In the following section, I will propose a formalized Feature-Checking process with which we can work on capturing these various behaviors of the *wh*-word constructions.

3. Feature-Checking Theory

In the first subsection of this section, I will briefly summarize the Feature-checking theory proposed by Chomsky (1992, 94), and in the second subsection formalize it as the Principle of Identity through modifying and clarifying Chomsky's argument.

3.1 Chomsky's Feature-Checking Theory

The Feature-checking Theory is expressed with the following statements by Chomsky (1992 and 1994).

(18) "... operations are driven by morphological necessity: certain features must be checked in the checking domain of a head, or the derivation will crash." (P. 35)

"The SPEC-head and head-head relations are therefore the core configurations of inflectional morphology." (P. 10)

“like verbs, nouns are drawn from the lexicon with all of their morphological features” (1994: 40).⁵

To interpret the Feature-checking theory with regard to functional categories, all the functional morphemes get their features checked in the domain of an appropriate functional category (head).⁶ For example, English *wh*-phrases get the feature [+WH] checked in the [SPEC, CP].

To speak more precisely in terms of MP, a functional category, as a checker playing a ‘mediating role’, has a set of features which motivate a functional morphemes’ movement, leading to check-off of the features possessed by functional morphemes through head-head adjunction or SPEC-Head agreement. For example, the sentence in (19a) *Mary loves John* has such a feature structure as in (19b) before movement.

- (19) a. Mary loves John.
 b. [AGRs T [Mary love John]]
 -PL_V +NOM +NOM -PL_V +ACC
 +3rd_V -PAST -PL_N +3rd_V
 -PL_N +3rd_N -PAST
 +3rd_N +ACC

⁵ In this paper, I use the terms morphological features, grammatical features, and functional features without distinction. There are roughly three kinds of functional features; syntactic features or categorial features expressed with [\pm N, \pm V], morphological features such as [\pm Nom, \pm Pl, \pm Past], and grammatical features such as [\pm FOC, \pm TOP]. I think a precise distinction of the last two is hard since in some languages topicalization is expressed with special morphemes while in others it is expressed with syntactic movement. Since the distinction is not crucial in this paper I will continue to use them without distinction.

⁶ One thing to be clarified in this context is the concept of ‘functional morpheme’ and that of ‘functional category.’ For example, Tense has a set of features with which it checks the features of verb and noun which raise later. That is, a Tense head is selected from the lexicon with just a set of features without any functional morpheme. In this case there are two objects which can be called a functional category: the Tense head and the Tense morpheme. In order to make our following discussion explicit, I will call the functional head possessing a set of functional features ‘functional category’, and the traditional functional category such as *that* or *not* ‘functional morpheme.’

To satisfy the Principle of Full Interpretation (FI), as stated in (20), the Checking Theory requires strong features to be checked off before SPELL-OUT (SO), since they are visible at PF and hence are not legitimate objects of PF. If a strong feature remains at PF, it cannot receive an interpretation, hence the derivation crashes, violating FI. Weak features do not create a problem at PF since they are invisible. Weak features wait until LF by Procrastinate (21), and they are checked there, satisfying FI again.

In the structure (19b), only the feature [+NOM] of T is strong and must be checked off before SO. According to Chomsky, before SO T adjoins to AGRs and forces *Mary* to substitute for [SPEC, AGRsP] in order to discharge the feature it shares with T, that is [+NOM]. Since all the other features are weak, they must wait until LF.

(20) **Full Interpretation**

Every symbol must receive an external interpretation by language-independent rules. (Chomsky 1992: 45)

(21) **Procrastinate**

LF movement is cheaper than overt movement. (Chomsky 1992: 42)

3.2 The Principle of Identity

In the preceding section I mentioned that features of functional categories motivate functional morphemes' movement. For example, verb-features of AGRs are ϕ -features (ϕ_V -features of AGRs) for verbs which adjoin to AGRs and its NP-features are ϕ -features (ϕ_N -features of AGRs) for NPs which substitute for

specifier-positions of AGRsP. These features come along with projections of AGRs from the lexicon and in the syntax they function to motivate movement. They are supposed to be checked off at a position of SPEC-Head agreement or a head-head adjunction when they meet a verb or an NP which has the same features.

To formalize the process of feature-checking, I introduce inverse features which are supposed to be a property of functional categories. That is, while the ordinary value $[\pm F]$ designates the morphological features of substantives or functional morphemes, the inverse value $[\pm F^-]$ does those of functional categories (nodes), as stated in (22). The introducing of inverse feature has at least two advantages. First, functional categories and functional morphemes can be defined distinctively in a more systematic fashion. That is, functional categories are checkers which have only inverse features, while functional morphemes are checkees which have ordinary features. Second, the feature-checking process can be more concretely visualized. If the concept of 'morphemes being licensed through feature-checking' is interpreted as 'getting their identities,' it is nicely expressed at the point where inverse features meet their ordinary counterparts. With this feature notation, the feature-checking process is formalized as the Principle of Identity (PI), as stated in (23).

- (22) a. $[\pm F]$ for functional morphemes
 b. $[\pm F^-]$ for functional categories

Accusative Case, which is already expressed in Chomsky's system.⁷ With this feature-notation, the above sample (24) can be expressed as in (25).

(25)	[AGRs	T	AGRo	[Mary	loves	John]]
		-PL _V ⁻	+NOM _V ⁻	+ACC _V ⁻	+NOM _N	-PL _V		+ACC _N	
		+3rd _V ⁻	-PAST _V ⁻	+ACC _N ⁻	-PL _N	+3rd _V			
		-PL _N ⁻	+NOM _N ⁻		+3rd _N	-PAST _V			
		+3rd _N ⁻				+NOM _V			
						+ACC _V			

This suggestion makes the feature system of functional categories more symmetric. That is, AGRs, AGRo, and T have two sets of features which are checked off through SPEC-Head agreement.⁸ Consequently, we have a generalization of the feature-system as stated in (26).

- (26) The functional categories entering into a SPEC-Head agreements have two sets of inverse features, one for NPs (Specifiers) and the other for verbs (heads).

The pairness of functional features is important to characterize the pairness of *wh*-word and the Q-morpheme *-nunci* or *-nunka* in Korean. In Section 4.2, I will argue that Korean *wh*-words and the Q-morpheme *-nunci* or *-nunka* are in the relationship of SPEC-head agreement in CP.

⁷ The paired features of AGRo also solve the problem caused by French verb raising before SO. If AGRo does not have these features and only the verb has the feature [+ACC], the raised verb before SO must be lowered to check off the [+ACC] feature of the object at LF. To avoid the lowering, we can introduce the copy theory. However, in my new feature system this is not needed.

⁸ In Section 5, I will show that Korean and Japanese *wh*-words are identified in CP in the same fashion.

4. Question formation and Property of Comp

According to PI (23), all the functional morphemes must be identified by some corresponding functional categories. Question sentences have special functional morphemes which must be identified by some functional category. Following Bresnan's (1979) argument that CP is responsible for the types of sentences, I will compare three languages, English, Korean, and Japanese, with respect to question formation. I will claim that the different syntactic processes of question formation between these languages result from different properties of the functional category Comp between languages. In other words, application of the PI to the Comps which have different properties from each other results in the syntactic differences of question formation between languages.

4.1 Properties of English Comp

To explain the subject-aux inversion in (27a), I assign the feature [$+Q^-$] to the matrix Comp and the ordinary [$+Q$] to the auxiliary *do(es)*, as stated in (28). This designation of the features is necessary in order to comply with the feature-checking theory.

- (27) a. Does John love Mary?
 b. Who does John love?
 c. *Who John loves?

- (28) a. [$+Q^-$] for Comp_M for direct *yes-no* question
 b. [$+Q$] for Auxiliaries

Two movements are involved in sentence (27b); auxiliary and *wh*-movement. For the *wh*-movement the Comp needs [+WH] besides [+Q] and *who* [+WH], as stated in (29). Even though the interrogative feature [+Q] has been assumed for the question formation from the Standard Theory, its necessity may be questioned since it is generally accepted that the feature [+WH] induces interrogation. However the function of [+Q] is totally different from that of [+WH]. The former plays the role of direct question maker, while the latter triggers a movement of *wh*-word which always binds a variable with which it induces interrogation. This is certainly a universal phenomenon in human languages. At least in English there is no embedded *wh*-question in which [+Q] is used and there is no direct *wh*-question without [+Q].

9

- (29) a. [+WH, +Q] for Comp_M for direct *wh*-question
 b. [+WH] for *Wh*-words

(30a) is the indirect question corresponding to (27a). The first question I address here is whether this embedded question has an operator which binds a variable. If (30a) corresponds only to *yes-no* questions like (27a), we can say there is no operator involved for the same reason that there is no operator involved in (27a).¹⁰ The second question is how we can designate features for the subordinator

⁹ I am not sure how we can derive the following question with this functional feature setting. I just assume that English has a null interrogative morpheme which has the feature [+Q]. In order to get [+Q] checked off, it moves to C which has the inverse feature [+Q].

(i) Who loves John?
¹⁰ Quirk et al. (1985) describe "the *yes-no* clause is introduced by the subordinators *whether* or *if*" (p. 1053). Even though they are not clear about whether alternative interrogative clauses like (ia) can be introduced by the subordinator *if* as in (iib), if it is possible, the correspondency between

if. As direct *yes-no* questions have the interrogative feature [+Q], it is certain that *if* also has the same feature and the corresponding Comp_E has the inverse feature [+Q⁻]. There is another thing to be considered. If the verb *ask* is subcategorized only for the Comp_E which has [+Q⁻], (30b) must be ruled in contrary to our expectations. In order for the Comp_E to be identified as the checker for the subordinator *if*, I suggest the feature [+SU] for it as its traditional name ‘interrogative subordinator’ implies.¹¹ That is, *if* has the features [+Q, +SU] and the Comp_E which is selected by the verb *ask* and checks off *if* has the inverse features [+Q⁻, +SU⁻], as given in (31). So, we have a new interpretation why (30b) is ruled out; it is because the feature [+SU⁻] is not checked off in (30b).

- (30) a. I asked if John loved Mary.
 b. *I asked did John love Mary.
 c. I asked who John loved.
 d. I asked whether John loved Mary.

- (31) a. [+Q, +SU] for *if*
 b. [+Q⁻, +SU⁻] for Comp_E

This argument is supported by the fact that *if* cannot appear in sentential subject constructions, as shown in (32). In our terms, (32) is ungrammatical because the feature [+SU] cannot be checked off because there is no Comp which has the inverse feature [+SU⁻].

(ia) and (ib) or (iia) and (iib) does not guarantee that (ia) does not have any operator. But I just assume it is not possible.

- (i) a. Are the banks open, or not?
 b. I wonder whether the banks are open.
 (ii) a. Are the banks open?
 b. I wonder if the banks are open.

Third, *wh*-words do not overtly move to [SPEC, CP] in Korean. And finally, in many cases, Korean embedded *wh*-words require a certain special interrogative morpheme on the Comp_E head position.

- (34) a. Jihi-ka cip-ey ka-nya?
 Jihi-N home-to go-Int
 'Is Jihi going home?'
- b. Jihi-ka encey cip-ey ka-nya?
 Jihi-N when home-to go-Int
 'When is Jihi going home?'
- c. Jihi-ka encey cip-ey ka-ess-nunci mwul-ess-ta.
 Jihi-N when home-to go-Past-Int ask-Past-Dec
 '(I) asked when Jihi went home.'

How can the first difference be explained in the feature-checking theory? For the first difference, I suggest the feature sets in (35). The only difference is that English Comp has the strong feature while Korean has the weak one. This is why in Korean there is no overt movement. In Korean, which is a head-final language, the word order does not vary depending up on the strength of interrogative feature. However, Procrastinate urges the interrogative feature to be weak in Korean, that is [+Q⁻].¹⁴

- (35) a. [+Q⁻] for Comp_M
 b. [+Q] for *-nya*

¹⁴ In other words, the phonetic forms of the sentences in (33) does not vary even when the interrogative feature is strong and checked off before SPELL OUT. However, when it is strong it must move before LF, and violates the principle of Procrastinate "LF movement is cheaper than overt movement" (= movement before SPELL OUT).

In the derivational theory of grammar, the second difference brings up a touchy problem; what is the categorial status of *-nya*? Because it never appears separated from verbs, it is definitely not an auxiliary as English *do* or *be*. There have been three analyses of the categorial status of the Korean sentential endings including the interrogative morpheme *-nya* in the literature. First, following Rizzi's (1990) argument that there possibly exists SPEC-Head agreement between *wh*-words and Comp, Jung (1992) treats it as a Comp, arguing that in Korean "a *wh*-phrase can obtain scope interpretation only if it is in a configuration of SPEC-Head agreement with the (interrogative) head Comp in its minimal domain" (p. 56). She proposes a double CP system in order to explain how two Comps can appear in (36a); *-nya* and *-ko* which is a generally accepted Comp. I do not agree with her on three points. First, we can license the interrogative morpheme in the single CP system if the morpheme is selected as a form conglomerated to verbs as in Chomsky's (1992) system. Second, the interrogative morpheme does not play the role of subordinator that is one of the major roles of complementizer. Third and most importantly, Korean has some cases in which embedded *wh*-words do not need an interrogative morpheme in Comp like English embedded *wh*-questions, as illustrated in (36b). This example weakens her argument that the interrogative morpheme is a Comp and that the other sentential endings *-ta* (declarative ending), *-la* (imperative ending), or *-ca* (cohortative ending) are also Comps which show complementary distribution with the interrogative morpheme, as shown in (36c).

- (36) a. Minho-ka Jihi-ka encey ka-nya-ko mwul-ess-ta.
 Minho-N Jihi-N when come-Int-Sub ask-Pa-Dec
 'Minho asked when Jihi would leave.'
- b. Minho-ka Jihi-ka encey o-n-ta-ko pokoha-ess-ta.
 Minho-N Jihi-N when come-Pr-Dec-Sub report-Pa-Dec
 'Minho reported when Jihi would come.'
- c. Jihi-ka o-n- -nya / *la-nya / *ca-nya?
 Jihi-N come-Pr- -Int Imp-Int Coh-Int
 'Is Jihi coming?' (intended reading)

Another pervasive view on the categorial status of sentential endings including the interrogative morpheme *-nya* is that they are modals projecting a Modal or Mood phrase between CP and IP (Ahn and Yoon 1989, Whitman 1989). This argument starts with the observation that the sentential endings including *-nya* co-occur with the generally accepted Comp *-ko* which is used as a subordinator as we can see in (36a & b).¹⁵ This is a quite attractive view in the derivational framework before the MP. However, in MP which allows dependent functional morphemes to come into the syntax as completely conglomerated forms with their substantives directly from the lexicon, introducing another functional category burdens the theory of grammar.¹⁶

I claim that the sentential endings including the interrogative morpheme *-nya* are just verb particles which come along with verb stems from the lexicon. This is a faithful application of the spirit of MP. As English verbs come along with all the inflectional

¹⁵ The subordinator *-ko* is optionally used. That is, the sentences in (23) are grammatical even if they do not have *-ko*.

¹⁶ Another largely accepted view is that the interrogative morpheme is base-generated in INFL (Kim 1991). However, as the original properties IP are composed of the phi-features and Tense, I do not accept this suggestion.

morphemes from the lexicon, Korean verbs are selected in the exactly same fashion from the lexicon. That is, verbs which have the interrogative morpheme have the feature [+Q] and the Comp has its inverse feature [+Q⁻]. The only difference from English is the strength of the feature; it is strong in English but weak in Korean. Notice that if the sentential endings, *-ta*, *-la*, *-ca*, and *-nya*, are not verb particles and they are introduced into the Comp node directly from the lexicon, we need another Comp node for them besides the one for the subordinator *-ko* and Jung's (1990) double CP analysis maybe plausible. As I claimed before, if they are verb particles, the double CP system is, however, no more necessary. They are base-generated into a verb node with a verb stem and at LF the entire verb, to get its feature checked, raises to Comp node where the subordinator *-ko* is already base-generated.¹⁷ I assume that this LF movement does not incur any structural problems.

Let us turn to the third difference. To explain this difference, I just follow Huang's (1982) proposal that the *wh*-words move to [SPEC, Comp] at LF in languages whose *wh*-words are in-situ at SS. In terms of the MP, *wh*-words have the feature [+WH] and

¹⁷ We could say that the subordinator *ko* is also affixed to verbs like the sentential endings in the lexicon since it always appears as a form affixed to a verb in the syntax. However, I reject this possibility for two reasons. First, the sentential endings determine semantic properties of verbs or sentences, but the subordinator *ko* does not. *ko* functions as a mere subordinator. Second, the appearance of *ko* is completely predictable in sentences, but the sentential endings are not as shown below;

- (i) Jihi-ka Minswu-eykey cip-ey ka-() () ha-n-ta.
 Jihi-N Minswu-to home-to go say-Pres-Dec
 'Jihi asks if Minswu goes home.'
 'Jihi suggests Minswu to go home together.'
 'Jihi orders Minswu to go home.'

In the second blank, only *ko* can be inserted. However, we cannot predict what will be inserted in the first blank. *-nya* (interrogative), *-ca* (cohortative), or *-la* (imperative) can be selected in the first blank.

respectively. Whereas, Korean has weak ones. The property of *wh*-words are the same in both languages; they have [+WH]. In Korean the interrogative feature [+Q] is possessed by the interrogative morpheme *-nya* which comes with verbs as a verb particle while in English it is possessed by auxiliary verbs.¹⁹ The embedded Comp which identifies the interrogative morpheme *-nunci* has the feature set [+WH_N⁻, +WH_V⁻].

A summary of the Properties of Comps is presented in (38).

(38) **Summary: Properties of Comp**

	English	Korean
Comp _M for <i>wh</i> -	[+ <u>WH</u> ⁻ , + <u>Q</u> ⁻]	[+WH ⁻ , +Q ⁻]
Comp _M for <i>yes-no</i>	[+ <u>Q</u> ⁻]	[+Q ⁻]
Comp _E for <i>wh</i> - (in the case of <i>ask</i>)	[+ <u>WH</u> ⁻]	[+WH _N ⁻ , +WH _V ⁻]
Comp _E for <i>yes-no</i>	[+ <u>Q</u> ⁻ , +SU ⁻] (for <i>if</i>)	[+Q ⁻ , +SU ⁻] (for <i>-nya-ko</i>)

4.3 Exploration of Comp Properties

The exceptions we discussed in Section 2 are classified into two classes. The first ones are the exceptions caused by the different types of interrogative morphemes; *-nunka*, *-nunci*, *-nya*. The second ones are the exceptions caused by different types of matrix verbs.

4.3.1 Classification of Interrogative Morphemes

First, examine how the interrogative morphemes are classified within the feature-checking theory. As illustrated in (39), the Q-morpheme *-nunci* is used only

¹⁹ The interrogative morpheme *-nya* has other properties, and there are several interrogative

the Comp has the weak inverse feature [+WH⁻] in Korean. The *wh*-words satisfy the PI at LF through SPEC-Head agreement in CP.

The fourth difference is that in Korean embedded *wh*-words, in many cases, require an interrogative morpheme on the head position of the embedded Comp. To explain this very difference, the *wh*-Constraints in (5) are proposed. A problem is, as we discussed in Section 2, in Korean embedded *wh*-words are allowed without an interrogative morpheme in many cases. To solve this problem, I suggest that these verbs are subcategorized for the Comps which do not have a feature for Q-morphemes and propose the feature in (37a) since the Q-morpheme *-nunci* always requires a *wh*-word (but not vice versa) and the features in (37b) for the corresponding Comp_E.¹⁸ These paired features are an expression of SPEC-Head agreement between *wh*-words and the Q-morpheme. In Section 4.3, I will discuss the *wh*-island effect of this SPEC-Head agreement.

- (37) a. [+WH_V] for *-nunci*
 b. [+WH_N⁻, +WH_V⁻] for Comp_E for *wh*-...-*nunci*

To sum up, the Comp properties of Korean direct questions are the same as those of English. The only difference is the strength of features. English Comp has strong features [+WH⁻, +Q⁻] and [+Q⁻] for *wh*-questions and *yes-no* questions,

¹⁸ In (i) *-nunci* is used without a *wh*-word. This sentence is interpreted as the same as English gloss. Therefore, I assume that the sentence in (i) has a null *wh*-word which corresponds to English *whether*.

(i) Jihi-ka pap-lul mek-ess-nunci all-n-ta.
 Jihi-N rice-A eat-Past-Int know-Pres-Dec

for indirect questions. This behavior of *-nunci* implies that there is a sort of Q-morpheme whose task is only to be paired with *wh*-words. The features in (37a) can explain why it cannot be used in direct questions like (39a). That is, there is no matrix Comp. which has the feature [+WH_V].

- (39) a. *Jihi-ka encey o-ess-nunci?
 Jihi-N when come-Pa-Int
 ‘When did Jihi come?’ (intended reading)
- b. Jihi-ka encey o-ess-nunci mwul / al / kiekha -ess-ta.
 Jihi-N when come-Pa-Int ask / know / remember -Pr-Dec
 ‘(I) asked / knew / remembered when Jihi went away.’

We can find similar phenomena in other languages. Rizzi (1990) presents the following examples from Kinande which show a one-to-one relationship between *wh*-words and subordinators.²⁰

- (40) a. IyondI y0 kambale alangIra
 who (cl. 1) that (cl. 1) Kambale saw
- b. BahI Bo kambale alangIra
 who (cl. 2) that (cl. 2) Kambale saw
- c. EkIhI ky0 kambale alangIra
 what (cl. 7) that (cl. 7) Kambale saw
- d. EBhI By0 kambale alangIra
 what (cl. 8) that (cl. 8) Kambale saw

The Q-morpheme *-nunka* is used to introduce indirect questions as well as direct ones, as illustrated in (41). When it is used in embedded clauses, it functions

²⁰ Jung (1991) cites these examples to support Kim’s (1991) generalization. I agree with them on

exactly the same way as *-nunci* does. Therefore, I claim that Korean has two kinds of *-nunka*, *-nunka₁* and *-nunka₂*, as shown in (42); one is used exactly as *-nunci*, and the other is used exactly as *-nya*.

- (41) a. Jihi-ka encey cip-ey ka-ess- **nunka**?
 Jihi-N when home-to go-Pa- Int
 'When did Jihi go home?'
- b. Jihi-ka encey ka-ess-**nunka** mwul / al / kiekha -ess-ta.
 Jihi-N when go-Pa-Int ask / know / remember -Pa-Dec
 '(I) asked / knew / remembered when Jihi went away.'
- (42) a. [+WH_V] for *-nunka₁* = *-nunci* (*-ka₁* in Japanese)
 b. [+Q] for *-nunka₂* = *-nya* (*-ka₂* in Japanese)

Japanese *-ka* is used for both questions as Korean *-nunka*, as shown in (43), and the distribution of this morpheme is exactly the same as the Korean *-nunka*. With this observation, I claim that Japanese has two kinds of *-ka*, as parenthesized in (42).

- (43) a. Dare-ga kuru **ka** ?
 who-N come Int
 'Who is coming?'
- b. Boku-wa dare-ga kuru **ka** sir-i-mas-en.
 I-Top who-N come Int know not
 'I do not know who is coming.'

Lastly, the Q-morpheme *-nya* is used for both questions as *-nunka*, but selected by only *ask*-type verbs. The feature [+Q] for *-nya* adopted in Section 4.2 is enough to explain this distribution.

- (44) a. Jihi-ka ka-ess-nya?
 Jihi-N go-Pa- Int
 ‘Did Jihi go away?’
- b. Jihi-ka encey ka-ess-nya-ko mwul / *al / *p okoha -ess-ta.
 Jihi-N when go-Pa- Int -Sub ask/ know / report -Pa-Dec
 ‘(I) asked/ knew / reported when Jihi went away.’

4.3.2 Properties of Comp_E Selected by Matrix Verbs

Let us now examine properties of Comps subcategorized by matrix verbs. The verbs we discussed so far are subcategorized into four types; *think*-type, *report*-type, *ask*-type, and imperative or cohortative verbs.

Let us start with the constructions of *think*-type verbs; *sayngkakha*- ‘think’, *al*- ‘know’, *kiekha*- ‘remember’, etc. The grammaticality of the sentences in (45a) says that these verbs select a Comp which identifies only *-ta* [+D] and *-ko* [+SU]. That is, the feature set of this Comp is as in (45). Whereas, the ungrammaticality of the sentences in (45b) which are formed by adding a *wh*-word *encey* ‘when’ to the sentences in (45a) says that *think*-type verbs are not subcategorized for the Comp whose feature set is [+WH_N⁻, +D⁻, +SU⁻] with which it can identify the *wh*-word. This line of thought is confirmed if we examine (45c). In these direct questions, the matrix Comp has the feature set of (47) and the *wh*-word *encey* [+WH_N] which could not be identified in (45b) can be saved. The fact that *wh*-words have only the matrix scope in these questions confirms the claim that the Comp selected by *think*-type verbs does not have the feature [+WH_N⁻] in this case.

- (45) a. Jihi-ka ka-n -ta / *-nya / *-la / *-ca -ko sayngkakha / al / kiekha -n-ta
 Jihi-N go-Pr-Dec/ Int/ Imp/ Coh -Sub think / know / remember -Pr-Dec
 ‘(I) think / know / remember that Jihi went away.’

- b. *Jihi-ka encey ka-ess-ta-ko sayngkakha / al / kiekha -n-ta.
 Jihi-N when go-Pa-Dec-Sub think / know / remember -Pr-Dec
 '(I) think / know / remember when Jihi went away.'
- c. Jihi-ka encey ka-ess-ta-ko sayngkakha / al / kiekha -ess-nya?
 Jihi-N when go-Pa-Dec-Sub think / know / remember -Pa-Int
 'When_i did you think / know / remember Jihi went away t_i?'
 not 'Did you think / know / remember when Jihi went away?'

(46) [+D⁻, +SU⁻]

As we discussed in the introduction, these verbs are included in the class for which Kim (1991) proposes the constraint (5). As predicted by the constraint, if the interrogative morpheme *-nunka₁* or *-nunci* is located in Comp, *wh*-words in the embedded sentences are allowed, as shown in (47).

- (47) a. Jihi-ka encey ka-ess-nunka₁ sayngkakha / al / kiekha -n-ta.
 Jihi-N when go-Pa-Int think / know / remember-Pr-Dec
 '(I) think / know / remember when Jihi went away.'
- b. Jihi-ka encey ka-ess-nunka₁ sayngkakha / al / kiekha -ess-nya?
 Jihi-N when go-Pa-Int think / know / remember-Pa-Int
 'Did you think / know / remember when Jihi went away?'
 not 'When_i did think / know / remember Jihi went away t_i?'

To represent this fact in our feature system, we can say that the embedded Comp has the feature set of (48) which identifies *-nunka₁* [+WH_V]. With this feature setting we can predict that *wh*-words bound by this Comp cannot cross over this Comp to take a wider scope. As we predict, the *wh*-word *encey* cannot take matrix scope in (47b), even if the matrix Comp has the feature set [+WH_N⁻, +Q⁻] as in the

sentences in (45c). That is, the scope interpretation can be predicted by the generalization (49).

(48) $[+WH_N^-, +WH_V^-]$

(49) Only the Comp which concerns SPEC-Head agreement exerts *wh*-island effect. (i.e., Comp of $[+WH_N^-, +WH_V^-]$ cf. (26))

To summarize, the *think*-type verbs are subcategorized for two Comps as stated in (50).

(50) *sayngkakha*:- ($\left\{ \begin{array}{l} [+WH_N^-, +WH_V^-] \\ [+D^-, +SU^-] \end{array} \right\}$)

Consider *report*-type verbs. This is one of the typical counterexamples against the *wh*-Constraint that *wh*-words must be governed by an interrogative morpheme in Korean. First, we need a Comp whose feature set is of (52a) in order to explain the sentences in (51a). This is the same Comp as for the *think*-type verbs that we discussed before.

- (51) a. Jihi-ka ka-ess-ta / *-nya / *-la / *-ca -ko **pokoha/ cenha/ alli** / -ess-ta.
 Jihi-N go-Pa-Dec- / Int/ Imp/ Coh -Sub report/ convey/ inform -Pa-Dec
 '(I) reported/ conveyed / informed that Jihi went away.'
- b. Jihi-ka **encey** ka-ess-ta-ko **pokoha/ cenha/ alli** / -ess-ta.
 Jihi-N when go-Pa-Dec-Sub report/ transfer/ inform -Pa-Dec
 '(I) reported/ conveyed / informed when Jihi went away.'

- c. Jihi-ka encey ka-ess-nunci pokoha/cenha/ alli / -ess-ta.
Int

As we examined, when a verb of this type is used in the matrix sentence, *wh*-words can take place in embedded sentences without any interrogative morpheme as in (51b), as well as in embedded sentences with an interrogative morpheme *-nunci* as in (51c). This dual property is captured by saying that these verbs are subcategorized for the two Comps that have the feature sets of (52b & c). The first one is selected in (51b) and identifies *encey* [+WH_N], *-ta* [+D], and *-ko* [+SU].²¹ The second one is selected in (51c) and identifies *encey* [+WH_N] and *-nunci* [+WH_V].

- (52) a. [+D⁻, +SU⁻] (cf. Comp_E for *think*-type verbs)
b. [+WH_N⁻, +D⁻, +SU⁻] ([(+WH_N⁻), +D⁻, +SU⁻])
c. [+WH_N⁻, +WH_V⁻]

With these feature sets, we can predict the scope interpretation of embedded *wh*-words. In the sentence in (53a), the *wh*-word *encey* can have both scopes; (52a) is selected for the matrix scope, and (52b) is selected for the narrow scope. Only the narrow scope reading is possible in (53b). In this case, (52c) is selected and the generalization (49) applies.

- (53) a. Jihi-ka encey o-n-Ta-ko pokoha-ess-nya?
Jihi-N when come-Pr-Dec-Sub report-Pa-Int
'(Did you) say when Jihi would come?'
'When (did you) say Jihi would come?'

- (56) a. [+Q⁻, +SU⁻] (cf. [+D⁻, +SU⁻] for *think-* or *report-*type verbs)
 b. [+WH_N⁻, +Q⁻, +SU⁻]
 c. [+WH_N⁻, +WH_V⁻]

The prediction of scope interpretation involved in these constructions is straightforward. The *wh*-word *encey* can have both scope interpretations in (57a); the Comp of [+WH_N⁻, +Q⁻, +SU⁻] is selected for the narrow scope, and the one of [+Q⁻, +SU⁻] for wide. In (57b), the *wh*-word *encey* can take only the narrow scope due to the Comp of [+WH_N⁻, +WH_V⁻]. Again, the generalization (49) applies.

- (57) a. Jihi-ka encey o-n-nya-ko mwul -ess-ny a?
 Jihi-N when come-Pr-Dec-Sub ask Pa-Int
 '(Did you) ask when Jihi would come?'
 'Wheni (did you) ask Jihi would come ti?'
 b. Jihi-ka encey o-nunci mwul -ess-nya?
 Jihi-N when come-Pr-Dec-Sub ask -Pa-Int
 '(Did you) say when Jihi would come?'

To summarize, the *ask*-type verbs are subcategorized as in (57).

- (58) *mwul*-(_____ { [+WH_N⁻, +WH_V⁻]
 { [(+WH_N⁻), +Q⁻, +SU⁻] })

Lastly, let us examine how imperative or cohortative verbs are subcategorized for Comps. As shown in (59a & b), imperative verbs allow only imperative sentential endings, and cohortative verbs allow only cohortative endings.

- (59) a. Jihi-ka cha-lul kochi- *nya/*ta/la/*ca-ko yokuha / myengha- ess-ta
 Jihi-N car-A fix -Int/ Dec/ Imp/ Coh-Sub require/ order -Pa-Dec
 'Jihi ordered (me) to fix the car.'

- b. Minho-lul manna *nya/ *ta/ * la / ca -ko ceyanha / ceuyuha-ess-ta.
 Minho-A meet -Int/ Dec/ Imp/ Coh-Sub suggest / propose -Pa-Dec
 '(I) suggested me to meet Minho.'

This means that these verbs are subcategorized only for the Comps of (60a) or (b), respectively.

- (60) a. [+IM⁻, +SU⁻]
 b. [+CO⁻, +SU⁻]

Interestingly, when the imperative verb *myengha-* 'order' occurs with subordinate *wh*-questions, it does not allow *wh... nunci*-constructions, contrary to the generalization proposed by Kim (1991). It rather allows only *wh...lako*-constructions as its subordinate clause, as shown in (61).

- (61) a. Jihi-ka mwues-lul kochi- la -ko yokuha / myengha- ess-ta
 Jihi-N what-A fix Imp-Sub require / order -Pa-Dec
 'Jihi ordered (me) what to fix.'
- b. *Jihi-ka mwues-lul kochi-nunci yokuha / myengha- ess-ta
 Jihi-N what-A fix Int require / order -Pa-Dec
 'Jihi ordered (me) what to fix.'

Similarly, the cohortative verb *ceyanha-* 'suggest' allows only *wh...lako*-constructions, but not *wh...nunci*-constructions, as shown in (62). With the same strategy, we can say that these verbs are subcategorized for the Comps that have the feature sets of (63 a) and (b), respectively.²²

- (62) a. Jihi-ka nwukwu-lul manna- ca -ko ceyanha-ess-ta.
 Jihi-N who-A meet - Coh-Sub suggest -Pa-Dec
 'Jihi suggested me whom to do.'

- b. ?*Jihi-ka nwukwu-lul manna- nunci ceyanha-ess-ta.
 Jihi-N who-A meet -Int suggest -Pa-Dec
 'Jihi suggested me whom to do.'

- (63) a. [+WH⁻, +IM⁻, +SU⁻]
 b. [+WH⁻, +CO⁻, +SU⁻]

With these feature sets, the scope interpretation of embedded *wh*-words is straightforward. In the sentence in (64a), the *wh*-word *encey* can have both scopes; the Comp of (63a) for the embedded scope, and the one of (60a) for the matrix scope. Similarly, in the sentences in (64b) the Comp of (63b) is selected for the narrow scope, while the one of (60b) for the matrix scope.

- (64) a. Jihi-ka mwues-lul kochi- la -ko yokuha / myengha- ess-nya?
 Jihi-N what-A fix -Imp-Sub require / order -Pa-Dec
 'Did Jihi order (y ou) what to fix?'
 'What did Jihi order (y ou) to fix?'
 b. Jihi-ka nwukwu-lul manna- ca -ko ceyanha-ess-nya?
 Jihi-N who-A meet -Coh-Sub suggest -Pa-Dec
 'Did Jihi suggest (y ou) whom to meet?'
 'Who did Jihi suggest (y ou) to meet?'

To summarize, the imperative and cohortative verbs are subcategorized as in

(65).

- (65) a. *myengha*-(_____ [(+WH⁻), +IM⁻, +SU⁻])
 b. *ceyanha*-(_____ [(+WH⁻), +CO⁻, +SU⁻])

To summarize this section, I pointed out four differences between English and Korean (and some of Japanese): (i) Usage of interrogative endings in Korean; (ii)

Categorial status of the sentential endings; (iii) No overt movement of *wh*-words in Korean; and (iv) the appearance of a special interrogative morpheme with *wh*-words in Korean. I tried to show that these differences are easily explained by using the Principle of Identity (the Feature-checking theory) with some newly introduced features.

I tried to show that the various behaviors of question formation depend on the difference of verb classes. In other words, different types of verb select different Comps, and the different properties of Comps induce the different behaviors of question formation.

I classified four types of verbs and showed that the traditionally accepted Constraint applies only to one class of verbs, that is, *think*-type verbs. The *report*-type verbs and *ask*-type verbs allow both constructions, *wh...-ko* and *wh...nunci*, allowing both scope interpretations of *wh*-words. In a completely opposite way to *think*-type verbs, imperative or cohortative verbs allow only the *wh...-ko* construction, allowing only narrow scope interpretation of *wh*-words. That is, the interrogative morpheme *-nunci* does not allow *wh*-words to take wide scope. This is analyzed by means of the SPEC-Head agreement phenomenon in CP. This is the typical difference from the actual interrogative morpheme *-nya*. All the differences are schematically expressed in our feature system, as shown in the following table. The optionally selected feature [+WH⁻] allows narrow scope reading of *wh*-words in the *wh...-ko* constructions.

(66)

CLASSIFICATION		FEATURES
Functional Morphemes		
Int. Morphemes	<i>-nunci</i>	[+WH _V]
	<i>-nunka₁</i>	[+WH _V]
	<i>-nunka₂</i>	[+Q]
	<i>-nya</i>	[+Q]
Sentential Endings	<i>-ta</i> (declarative)	[+D]
	<i>-la</i> (imperative)	[+IM]
	<i>-ca</i> (cohortative)	[+CO]
Subordinator	<i>-ko</i>	[+SU]
Wh-words	<i>encey</i> 'when' <i>nwukwu</i> 'who' <i>mwues</i> 'what' etc.	[+WH _N]
Functional Categories		
Matrix Comps	for <i>wh</i> -questions	[+WH _N ⁻ , +Q ⁻]
	for <i>yes-no</i> questions	[+Q ⁻]
	for non-questions	[+D ⁻], [+IM ⁻], or [+CO ⁻]
Embedded Comps	for <i>think</i> -type verbs	[+WH _N ⁻ , +WH _V ⁻] [+D ⁻ , +SU ⁻]
	for <i>report</i> -type verbs	[+WH _N ⁻ , +WH _V ⁻] [(+WH _N ⁻), +D ⁻ , +SU ⁻]
	for <i>ask</i> -type verbs	[+WH _N ⁻ , +WH _V ⁻] [(+WH _N ⁻), +Q ⁻ , +SU ⁻]
	for Imp/Coh verbs	[(+WH _N ⁻), +IM ⁻ , +SU ⁻] [(+WH _N ⁻), +CO ⁻ , +SU ⁻]

5. Conclusion

In this paper, I reconsidered *wh*-question formation in Korean, Japanese, and English, and proposed a new feature system for Comp in the framework of the MP.

It is claimed that the interrogative morpheme *-nunci* and *-nya* must be treated differently. I classified four types of verbs and showed that the *wh*-Constraint applies only to *think*-type verbs in Korean, and that this phenomenon must be and

successfully explained in the feature-checking theory. In addition, within the proposed feature system (SPEC-Head agreement) the scope interpretation of embedded *wh*-words is explained without the *wh*-Constraint. For the purpose of the study, I formalized the feature-checking theory as the Principle of Identity and clarified some morphological differences between English and Korean.

The result of the study shows us that Korean has a well-developed system of sentential endings which are tightly connected to the appropriate group of verbs. The *wh*-Constraint applies to Japanese consistently, but not to Korean. It seems that the tightness of the connection between the sentential endings and verbs overwhelms the Constraint.

References:

- Ahn, H.-D., and H.-J. Yoon (1988). Functional Categories in Korean. In S. Kuno et al., eds. *Harvard Studies in Korean Linguistics III*. Seoul: Hanshin Publishing Co.
- Bobaljik, J. (1993). On Ergativity and Ergative Unergatives. In Phillips, C. ed. *MIT Working Papers in Linguistics* 18: 45-88.
- Bresnan, J. (1979). *Theory of Complementation in English Syntax*. Garland Publishing Inc.
- Carnie, A. (1993). Nominal Predicates and Absolutive Case Marking in Irish. In Phillips, C. ed. *MIT Working Papers in Linguistics* 18: 89-132.
- Chomsky, N. (1982). *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris Publications.
- _____ (1988). Some Notes on Economy of Derivation and Representations. In Freidin, R. ed. *Principles and Parameters in Comparative Grammar*: 406-454. Cambridge, Massachusetts: MIT Press.
- _____ (1992). Minimalist Program for Linguistic Theory. *MIT Occasional Papers in Linguistics* 1
- _____ (1994). Bare Phrase Structure. *MIT Occasional Papers in Linguistics* 5.
- Collins, C. & H. Thrainsson (1993). Object Shift in Double Object Constructions. In Phillips, C. ed. *MIT Working Papers in Linguistics* 18: 134-174.
- Fukui, N. (1986). *A Theory of Category Projection and Its Applications*. Ph.D. Dissertation, MIT.
- Grimshaw, J. (1979). Complement Selection and the Lexicon. *LI* 10-2: 309-356.
- _____ (1991). Extended Projection. ms., Brandeis University.
- Han, K.-S. (1968). *Palun-Mal Kowun-Mal*. [Correct Language and Beautiful Language]. Seoul: Yungmoonsa.
- Huang, C.-T. (1982). *Logical Relations in Chinese and the Theory of Grammar*. Ph.D. Dissertation, MIT.
- Jung, Y.-J. (1992). *A Restrictive Theory of Functional Categories and Their Parametric Variation*. Ph.D. Dissertation, University of Washington.
- Kim, S. (1991). *Chain Scope and Quantification Structure*. Ph.D. Dissertation, Brandeis University.
- Lee, C.-M. and Y.-N. Ko (1987). *Enehak Sacen*. [Dictionary of Linguistics]. Seoul: Pakyoungsa.
- Nahm, K.-S., & Y.-K. Ko (1985). *Pyojun Kuke Munpepnon*. [A Standard Korean Grammar]. Seoul: Tap Press.
- Nishigauchi, T. (1990). *Quantification in the Theory of Grammar*. Dordrecht: Kluwer Academic.
- Pollock, J. (1989). Verb Movement, Universal Grammar, and the Structure of IP. *LI* 20: 395-424.
- Quirk, R. et al. (1985). *A Comprehensive Grammar of the English Language*.

A Case-Tier Approach to Finnish: Feature-based Definition of Case Domain

Yongkil Jeong

1. Introduction

It has long been argued in the GB theory that the subject (=Spec of IP) is assigned nominative case by finite INFL and the object is assigned accusative case by a transitive verb (Chomsky 1981, 1986a, and 1986b). However, languages such as Icelandic and Finnish which show various case alternations seem to contradict this argument. As an alternative to the GB case theory, Zaenen *et al.* (1985) and Yip *et al.* (1987) suggest the case-tier approach. The general assumption in the case-tier hypothesis is that the domains that supply a case tier are determined on the basis of categories. In English, for example, Ss (=IPs) form a case tier consisting of two surface cases, that is, nominative and accusative case, and PPs form a case tier containing only accusative case. Further, Maling (1992) suggests that IPs supply their own case-tier, whereas VP complements such as infinitives do not. Instead of the category-based definition of case domain, in this paper, we will suggest a feature-based definition of case domain. The basic idea is that not every IP but only IP with a [+AGR] feature can supply a case tier. Case alternations in Finnish can be accounted for in a principled way under the feature-based definition of case domain. In section 2, we will see case alternations between nominative and accusative case in Finnish, and in section 3, we will review a case-tier approach to Finnish (Maling 1992), which suggests a category-based definition of case domain. Finally, in section 4, we will suggest a feature-based definition of case domain within the case-tier hypothesis.

2. Case Alternation in Finnish

2.1 Nominative Object

Finnish is an accusative language so that the subject of a finite verb is assigned nominative case and the object of a finite verb receives accusative case as shown in (1).

(1) a. *mies tulee*
 nom. 3d.sg.
 'the man comes'

b. *mies saa kirjan*
 nom. 3d.sg. acc.
 'the man gets the book'

In some environments, however, the object of a transitive verb is assigned nominative, not accusative case. There are basically three environments in which nominative case is assigned to the object. First, nominative case is assigned to the object of an imperative, as illustrated in (2). In (2a) *tyttö* 'the girl' is the object of *saata* 'take' but is assigned nominative rather than accusative case. (2a) and (2b) show that no agreement exists between the verb and the nominative noun, which indicates that the noun is not the subject.¹

(2) a. *saata tyttö kotiin!*
 impv. nom. adv.
 '(you.sg.) take the girl home'

b. *saattakaa tyttö kotiin!*
 impv. nom. adv.
 '(you.pl.) take the girl home'

¹. See Timberlake (1975), which argues that the grammatical subject must agree with the verb.

The imperative in Finnish is usually limited to the second person singular and plural and first person plural forms. However, there is a third person optative form, where the object is assigned accusative case. Note here that the subject appears in the optative form and is assigned nominative case.

- (3) *han saattakoon tytön kotiin!*
 nom. opt.3d.sg acc. adv.
 'may he take the girl home'

Second, nominative case is assigned to the object of an impersonal passive as shown in (4). Here again we need to note that the nominative nouns in (4a) and (4b) cannot be the grammatical subject since there is no subject-verb agreement. The verbs in (4) show the same agreement whether the nominative is singular or plural.

- (4) a. *sinne viedaan lahja*
 adv. pass. nom.
 'the present will be taken there'

- b. *sinne viedaan lahjat*
 adv. pass. nom.pl.
 'the presents will be taken'

Finally, nominative case is assigned to the object of an infinitive when the infinitive is used as the subject of a matrix clause. It is quite clear that *se* 'it' and *kirje* 'a letter' in (5) are the object of the infinitive verbs even though they are assigned nominative case. The subjects of the infinitive may optionally be assigned genitive case.

- (5) a. *minun taytyy tehda se*
 gen. 3d.sg. inf.I nom.
 'it is necessary for me to do it'
- b. *hanen taytyy kirjoittaa kirje*
 gen. 3d.sg. inf.I nom.
 'it is necessary for him to write a letter'

Now let us consider more interesting cases. In Finnish, nominative case is also assigned to the object of an infinitive when it is used as an adjunct or a complement of any one of the above-mentioned three cases.

- (6) a. *kaske hanen tehda se!*
 impv. gen inf.I nom.
 'tell him to do it'
- b. *mene kaupunkiin ostamaan uusi hattu!*
 impv. ill. inf.III adj. nom.
 'go to town to buy a new hat'
- (7) a. *se kaskettiin tehda*
 nom. pass. inf.I
 'it was ordered to be done'
- b. *minua pyydettiin jattamaan tama kirje teille*
 part. pass. inf.III dem. nom. ill
 'I was-asked to-deliver this letter to-you'

(8) a. *hanen taytyy menna kaupunkiin ostamaan uusi hattu*
 gen. 3d.sg. inf.I ill. inf.III adj. nom.
 'she has to go to town to buy a new hat'

b. *minun taytyy kayda noutamassa kirje*
 gen. 3.sg. inf.I inf. III nom.
 'I have to go to get the letter'

In (6a), the infinitive is used as the complement of an imperative verb *kaske* and the object of the infinitive *se* takes nominative case. In (6b), the infinitive functions as an adjunct of an imperative verb *mene* and the object *hattu* is assigned nominative case. In (7), nominative case is assigned to the object of the embedded infinitive, which is used as the complement of an impersonal passive verb. Both (8a) and (8b) contain two infinitive verbs. In (8a), the most deeply embedded infinitive *ostamaan* is governed by another infinitive verb *menna*. The former is used as the adjunct of the latter. Here nominative case is assigned to the object of the most deeply embedded infinitive, that is, *hattu*. It is exactly the same in (8b).

2.2 Accusative Object

In this subsection, we will see the cases in which the object of an infinitive takes accusative case. First, accusative case is assigned to the object of an infinitive when the infinitive is followed by a finite verb, as illustrated in (9).

(9) a. *han antoi pojan tehda sen.*
 nom. 3.sg. gen. inf.I acc.
 'he had the boy do it'

b. *kavin noutamassa kirjeen postista*

1.sg inf.III acc. elat.

'I went to get the letter from the office'

In (9a), the embedded infinitive is used as the complement of a finite verb *antoi*, and in (9b), the infinitive is used as an adjunct of a finite verb *noutamassa*. In both cases, the object of the infinitive takes accusative case. These examples show that whether or not the object is assigned nominative case depends upon the context in which the infinitive is embedded. Now let us look at the sentence in (10).

(10) *kaske hanen ainakin koettaa tulla tekemaan se!*

impv. gen. adv. inf.I inf.I inf. III nom.

'tell him to at least try to come to do it!'

The condition for assigning case to the object of an infinitive seems to be a recursive one. This recursive property can even lead to a sentence in which a triply embedded infinitive takes a nominative object *se* in (10). Note that the Inf.III is ultimately embedded under an imperative verb *kaske*.

Second, infinitives that are inflected with the possessive affix for the logical subject take an accusative object regardless of the matrix environment.²

(11) a. *matkusta Suomeen oppiaksesi suomen kielen!*

impv. ill. inf.I+2.sg. gen. acc.

'travel to Finland to learn Finnish language'

² This kind of infinitive is called an inflected infinitive in Timberlake (1975).

b. *saadakseni rahan yritin kaikenlaista.*

inf.I+1.sg. acc. 1.sg. part.

'to get the money I tried everything'

In (11a), the infinitive is governed by a non-finite verb, that is, an imperative *matkusta*. Nonetheless, the object *kielen* is assigned accusative case contrary to the observation in 2.1. Here we can see that the infinitive verb is inflected for its logical subject. An agreement is found in number and person between the verb and its logical subject, even though the subject is not overtly realized. In (11b), on the other hand, the infinitive is governed by a finite verb *yritin*. As expected, the object takes accusative case.

Such case alternations as we have seen seem not to be accounted for by the traditional GB case theory.

3. Case Tier Approach to Finnish

3.1 Category-based Definition of Case Domain

The case-tier hypothesis was first proposed in Zaenen, Maling and Thrainsson (1985) and has been supported by Yip, Maling and Jackendoff (1987) (YMJ, hereafter) and Maling (1992), among others. The basic idea is that syntactic surface case forms a tier which is autonomous of phrase structure and that it is associated with NPs of the phrase structure by general principles. One of the basic principles is that case assignment applies only to items strictly within a domain. It is assumed that the relevant domains include S (=IP) and NP. The appropriate definition of domain for the case-tier hypothesis is given in YMJ (1987).

(12) a. A node Y is in the case domain of a node X iff Y is dominated by X, and X supplies a case tier C.

- b. A node Y is in the **strict case domain** of a node X iff Y is in the case domain of X , and there is no node Z such that X dominates Z and Y is in the case domain of Z .

The unmarked situation can be illustrated from English. A simple transitive sentence contains two NP's and two surface cases: N(ominative) and A(ccusative)³. These cases universally form a linearly ordered tier. The result of one-to-one, Left-to-Right association of an equal number of NPs and cases is illustrated in (13)⁴. The nominative case is associated with the leftmost NP, which is the subject and the accusative case with the remaining NP, that is, the object.

- (13) [_S John hit Mary].
 ↑ ↑
 N A

An intransitive sentence has the same case tier, but only one NP to receive case. Thus, the accusative is simply unrealized.

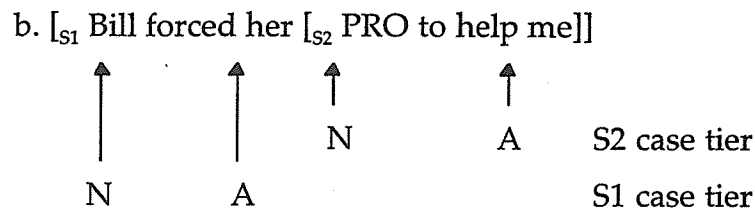
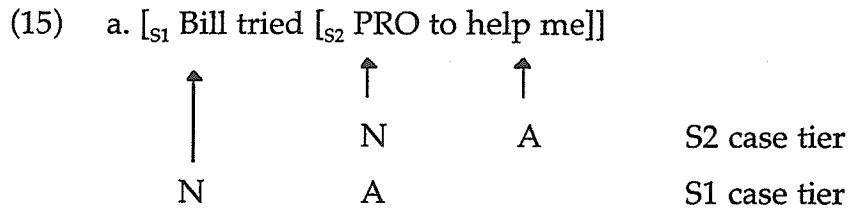
- (14) [_S John paused].
 ↑
 N A

The case tier hypothesis proposed by YMJ (1987) assumes that PRO receives structural case. Let us look at (15). There are two case domains in both (15a) and (15b). The matrix clause and its complement clause each have their own

³. They do not preclude the possibility of the presence of a VP (or V')-supplied tier in English. In other words, the IP-tier contains only a nominative case and the VP-tier contains only an accusative case. If this is so, it seems to be hard to find any significant difference between the GB case theory and the case-tier theory.

⁴. To be more precise, the linear order of NPs does not matter in the case-tier hypothesis. Rather, an abstract hierarchical order of grammatical functions is relevant. Thus the subject which is grammatically higher than the object is assigned nominative case.

case tier consisting of nominative and accusative cases, which are associated with NPs within them, including PRO.⁵



The nominative case of the S2 case domain goes to PRO and the accusative of the S2 goes to the object of the S2 in both (15a) and (15b). The nominative case of the matrix clause S1 goes to the matrix subject in both cases. There is no NP remaining in (15a) and thus the accusative case of the S1 is simply unrealized. In (15b), however, there is an NP in the S1 case domain, which is assigned the accusative case of the S1. According to the definition (12), case assignment in the matrix clause in (15a) cannot attach accusative case to any NP in its complement clause.

3.2 Maling's (1992) approach to Finnish Infinitives

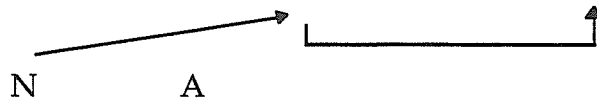
Maling (1992) provides a case-tier approach to Finnish. In order to account for the nominative object of an infinitive, she follows Vainikka (1989), which differentiates the categories of finite clauses and non-finite clauses.

According to Vainikka (1989), there are two kinds of verbal complements: IP and VP (or V'). Tensed clauses are IPs and non-finite clauses

⁵. It is against the GB case theory that PRO is assigned a structural case. Note that PRO is assumed to receive a Null Case in the Minimalist Program (Chomsky 1992).

such as infinitives are VP or V'⁶. The non-finite complements lack the projections of INFL. Adapting Vainikka's analysis of infinitival complements to the case-tier hypothesis, Maling (1992) assumes that VP-complements are transparent to case assignment from the matrix clause. In other words, VPs or V's are not a barrier to case-assignment. On the other hand, finite complements are assumed to have their own case tier. They are not transparent to case assignment from the matrix clause. The structural cases can percolate down into the infinitival VP-complement to any receptive NP. Now let us look at how Maling's (1992) analysis works.

(16) a. [_{IP} mene kaupunkiin [_{V'} ostamaan uusi hattuu]]! (6b)



b. [_{IP} han antoi [_{VP} pojan tehda sen]]. (9a)

(16a) has only one case domain, i.e., the matrix IP domain. The V' complement does not form its own case domain. There is no candidate NP for a structural case in the matrix clause. Therefore, nominative case goes down into the embedded clause, looking for an NP and finally it goes to the only NP *hattuu*. In (16b), only the matrix IP can supply its case domain. The VP complement cannot form a case tier. Here the subject of the matrix IP takes

⁶ Vainikka (1989) assumes that the category of inf.I (or TA-inf.) is VP and that of inf.III (or MA-inf.) is V'. The TA-infinitive has the distribution of a direct object, while the MA-infinitive has the syntactic category and distribution of a locative PP.

⁷ G (genitive case) is not a structural case in this analysis. It is one of a set of lexical cases, which are assigned independently of the structural cases.

the first structural case, i.e., nominative case. The remaining structural case, i.e., accusative case cannot find a candidate for itself within the matrix IP and therefore it goes down into the VP complement. Finally the object in the complement takes accusative case. Maling's analysis works in the same way in the cases where infinitives are governed by an impersonal passive or an infinitive as in (7) and (8).⁸

Maling (1992) does not mention the inflected infinitive in Finnish. It is not clear how her analysis would work in case of inflected infinitives. She does not mention whether PRO is assigned a structural case or not. One possible structure of (11a) would be as in (17).

- (17) [_{IP} matkusta Suomeen [_{VP} oppiaksesi suomen kielen]]! (11a)
 impv. ill. inf.I+2.sg. gen. acc.
 'travel to Finland to learn Finnish language'

According to Maling (1992), there is only one case domain in (17) since the embedded VP cannot form a case tier. Nominative case does not find any NP in the matrix. Therefore it goes down into the embedded VP and finally settles down in the object of the infinitive. This results in a wrong prediction. Another possibility is to put a PRO or a *pro* in the embedded infinitive in (17).⁹

- (18) [_{IP} matkusta Suomeen [_{VP} PRO oppiaksesi suomen kielen]]! (11a)
 impv. ill. inf.I+2.sg. gen. acc.
 'travel to Finland to learn Finnish language'

⁸ It is assumed that the matrix clause supplies its own case domain whatever its type may be.

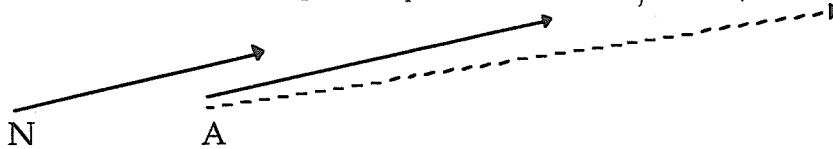
⁹ There may be another alternative. One can argue that the inflected infinitive is actually IP rather than VP or V'. If so, the infinitive supplies its own case domain as follows;

- (i) [_{IP1} matkusta Suomeen [_{IP2} PRO oppiaksesi suomen kielen]]!

Here the nominative case of IP2 is associated with PRO and the accusative case of IP2 with *kielen*, which is a correct prediction.

Here the nominative case of the matrix domain is associated with the subject PRO and the accusative case with the object *kielen*. This results in a correct prediction. We need the assumption that PRO is assigned a structural case. Once we assume that PRO is assigned a structural case, however, we cannot account for the case assignment in the VP-complement in (10).

(19) [_{IP} kaske hanen ainakin [_{VP} PRO_i koettaa [_{VP} PRO_j tulla [_V tekemaan se]]]]! (10)



The nominative case of the matrix case-tier should be associated with PRO_i and the accusative with PRO_j. If so, the accusative should spread to the object of the most deeply-embedded infinitive.¹⁰ However, this is not the case¹¹.

4. Feature-based definition of case domain

In this section, I will suggest a feature-based definition of case domain within the case-tier hypothesis. YMJ (1987) argues that every IP forms a case-tier. Maling (1992) also argues that IPs always form their own case-tier, whereas VPs do not. In the traditional GB case theory, finite INFL but not non-finite INFL assigns nominative case to its specifier position and a transitive verb assigns accusative case to its complement. Adapting nominative case assignment in the GB theory to the case-tier theory, I suggest that only finite IPs can supply their own case-tier. The case domain in this system I am proposing is determined based upon features of INFL, which is different from Maling (1992) and YMJ (1987), which assume that case domain is determined based upon categories. Case alternations in Finnish seem to be better

¹⁰. In (19), there is no structural case available for the object of the most deeply embedded infinitive after both nominative and accusative are already assigned to other NPs. In this case, the rightmost structural case spreads to the remaining NPs rightwardly.

¹¹. It seems to be possible to assume that the covert subject of an inflected infinitive is a *pro* and that of a non-inflected infinitive is a PRO and that *pro* may be assigned case whereas PRO may not. As we have seen in (15), however, PRO must be assigned case in English.

accounted for by the feature-based definition than the category-based definition of case domain. Here I am proposing that the case-tier is licenced by INFL with a [+AGR] feature¹². The maximal projection (=IP) of an INFL with an agreement feature is itself a case domain. On the other hand, the maximal projection of an INFL without an agreement feature cannot be a case domain by itself and must be part of the case domain of its matrix IP.

It is generally assumed that INFL consists of two features, that is, AGR and Tense. Therefore we may divide clauses in Finnish into four possible types, depending upon the feature combination of their INFL.

(20) Clausal Types in Finnish

- finite tensed clauses: [+AGR +Tense]
 inflected infinitives: [+AGR -Tense]
 impersonal passives: [-AGR¹³ +Tense]
 infinitive (& impv.¹⁴): [-AGR -Tense]

Now let us see how this feature-based definition of case domain works in Finnish. In the cases of finite indicative active sentences, IPs supply their own case-tier consisting of two surface cases, which are mapped into NPs one-to-one and left-to-right.

¹². Alternatively, it can be argued that INFL with [+AGR] or [+Tense] licenses a case-tier. If impersonal passive complements do not supply their own case-tier since the head lacks agreement, the following example seems not to be accounted for.

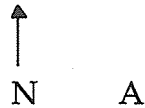
- (i) luulen etta mies kutsutaan.
 1st.sg. comp. nom. pass.
 I think that the man will be invited

I assume that the head COMP licenses a case domain in the embedded clause since it blocks case transmission from the matrix into the embedded clause.

¹³. In impersonal constructions, the verb always shows the default agreement, that is, 3rd singular. This default agreement is considered [-AGR].

¹⁴. It seems that imperative verbs in Finnish show agreement with the subject, whether it is overt or not. However, I assume that there is no agreement in Finnish imperatives. See Timberlake (1975:pp 219-220), which argues that although the logical subject of imperatives is predictable from the verb agreement, it does not arise through a free choice of possible subjects in the way that the subject of a finite indicative active verb does.

(21) a. [_{IP} mies tulee] (1a)



b. [_{IP} mies saa kirjan] (1b)



We have seen in section 2 that whether or not the object of an infinitive is assigned nominative case depends upon the context in which the infinitive is embedded. First, consider (5a), which would have the following D-structure¹⁵.

(22) [_{IP} e taytyy [_{IP} minun tehda se]] (5a)

3d.sg. gen. inf.I nom.

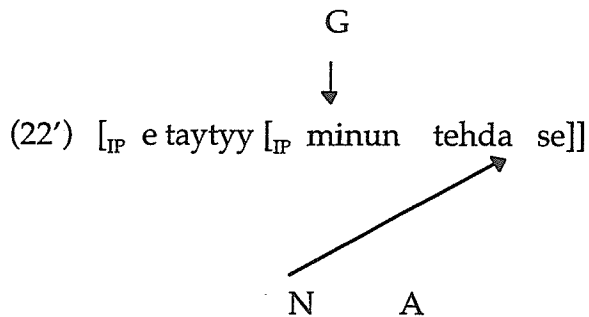
'it is necessary for me to do it'

In (22), the embedded IP cannot supply its own case-tier since the head INFL lacks an agreement feature. Therefore, there is only one case-tier in (22), which is supplied by the matrix IP. The subject of the embedded IP is assigned genitive case, which is a lexical case. There remains only one candidate for the structural cases, that is, the object of the embedded IP¹⁶. The object is assigned nominative case by the principles of the case-tier hypothesis.

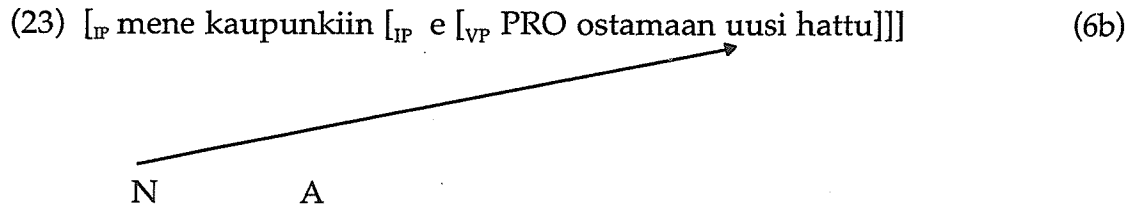
¹⁵. Alternatively, we may have the following D-structure for (5a).

(i) [_{IP} [_{IP} minun tehda se] taytyy]

¹⁶. Sentential arguments are not dealt with in this paper.



Now let us consider (6b) in which the infinitive is the complement of a non-finite verb and (9a) in which the infinitive is the complement of a finite verb.



(23) has only one case-tier since the embedded IP cannot form a case-tier. Nominative case is mapped into the embedded object since there is no NP in the matrix. However, there arise two questions. One is why PRO in the embedded VP is not assigned any structural case and the other is why PRO does not raise to [SPEC, IP], thus remaining in [SPEC, VP]. We assume that the subject-raising from [SPEC, VP] to [SPEC, IP] is triggered by the feature [+AGR]. The head INFL of infinitives is classified as [-AGR, -Tense]. Therefore, there seems to be no reason to raise PRO to [SPEC, IP]¹⁷. This is the answer to the second question. For the first question, I assume that [SPEC, IP] is the candidate position for structural cases and [SPEC, VP] is the candidate position for lexical cases¹⁸. There still remains another question. How can the PRO in [SPEC, VP] be assigned case¹⁹? In (23), no lexical case is assigned to PRO²⁰. It

¹⁷. For the accusative object of an infinitive in English, we have to stipulate that PRO in [SPEC, VP] always raises to [SPEC, IP], which seems ad hoc. Perhaps the lexical head 'to' is responsible for the obligatory raising of PRO.

¹⁸. In the case-tier theory, structural cases like nominative and accusative do not have anything to do with syntactic positions. However, it is assumed that VP is the semantic projection of a verb.

¹⁹. Following YMJ (1987), I assume that PRO must have a case.

²⁰. It is not clear whether a lexical case can be assigned to PRO. This is not relevant to our discussion.

cannot be assigned accusative since the assignment of accusative crosses the assignment of nominative, which is not possible in the case-tier hypothesis. However, there is a way for PRO to be assigned case, that is, case-spreading. Interestingly, the spreading is leftward in this case²¹.

(23') a. [_{IP} mene kaupunkiin [_{IP} e [_{VP} PRO ostamaan uusi hattu]]]



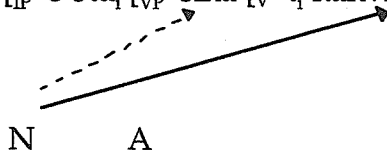
This leftward case-spreading is supported by imperatives in Finnish. There are imperative constructions in which the subject and the object are both assigned nominative case although the subject must appear postverbally.

(24) a. ota sina kahvi kaapista!
 impv. nom. nom. elat.
 you take the coffee from the cupboard

b. *ota sina kahvin kaapista!
 impv. nom. acc. elat.

The structure of (24a) would be like (25). Here, the subject is not assigned any lexical case at the level of direct case-mapping.

(25) [_{IP} e ota_i [_{VP} sina [_V t_i kahvi kaapista]]]

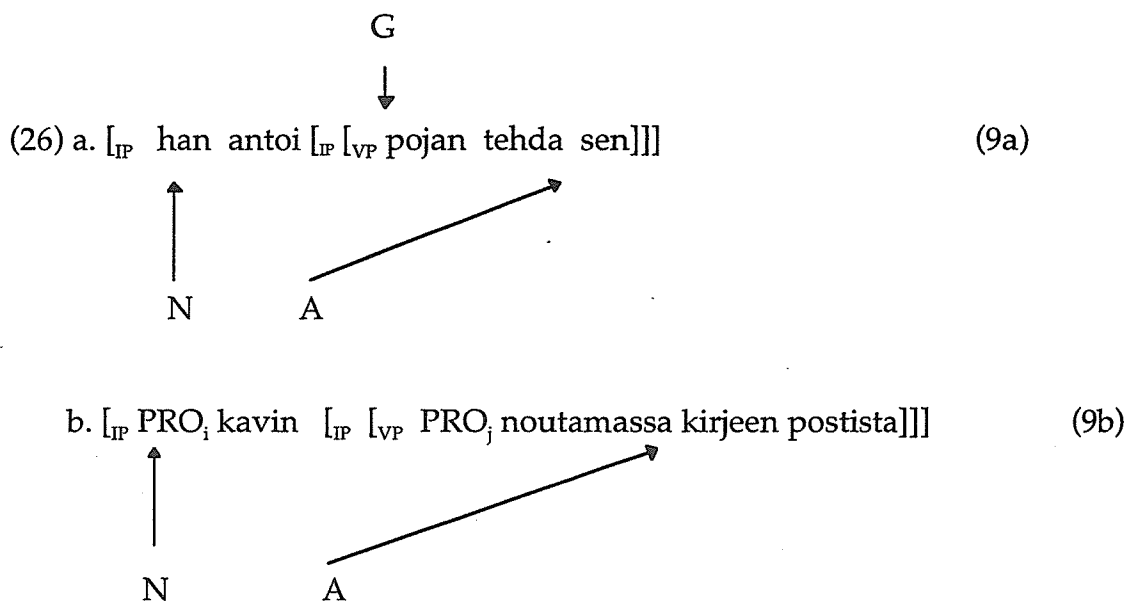


²⁰. It is not clear whether a lexical case can be assigned to PRO. This is not relevant to our discussion.

²¹. We assume that [SPEC, VP] is open to case spreading, but not to direct mapping of structural cases.

The nominative case is mapped into the object since the subject remains in [SPEC, VP] position. After the nominative case is assigned to the object, however, it spreads leftward to the subject. We assume three stages of case assignment: (i) lexical case assignment at DS, (ii) structural case assignment at SS, and (iii) case-spreading²².

Now let us look at infinitives that are embedded within a finite tensed clause.

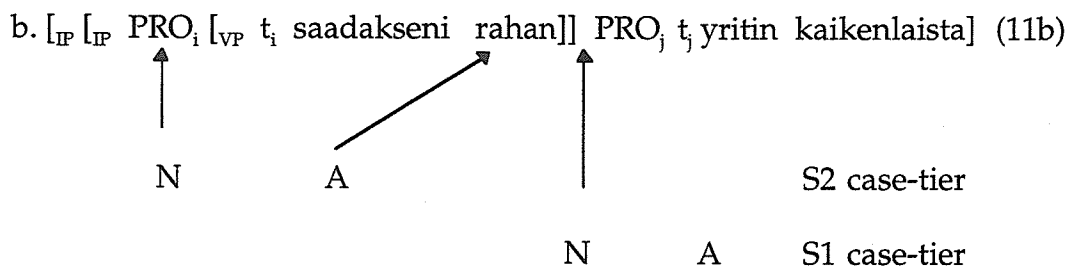
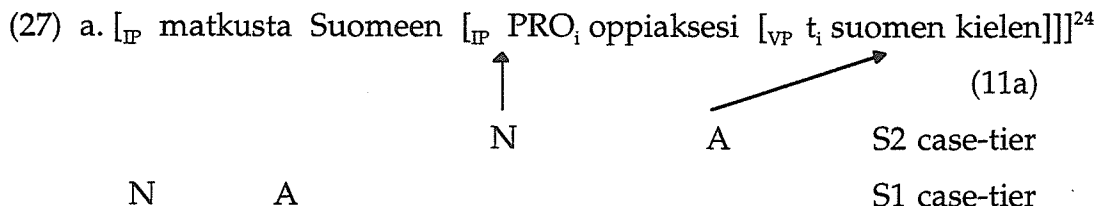


In (26a, b), the embedded IPs do not supply their own case-tier since their head INFL lacks the agreement feature. Thus, only one case-tier is available. In (26a), the nominative and the accusative case are mapped into the subject of the matrix clause and the object of the embedded clause, respectively. In (26b), however, there are three NPs competing for structural cases: PRO_i, PRO_j and *kirjeen*. Under the analysis presented here, nominative case is assigned to the PRO_i, which raises from [SPEC, VP] to [SPEC, IP]. The agreement between PRO_i and the verb *kavin* triggers the raising of PRO_i. PRO_j cannot be assigned any syntactic case since it is in [SPEC, VP]. There is no reason to raise it to [SPEC, IP] since there is no agreement between PRO_j and *noutamassa*. Therefore, the

²² It is not clear at which level case-spreading occurs. It may occur at PF and thus the result of case-spreading is like default case.

accusative case is assigned to the object of the infinitive *kirjeen*. After syntactic case-assignment, the accusative case of *kirjeen* spreads to PRO_i ²³.

Finally, let us see how the feature-based definition of case domain accounts for inflected infinitives whose objects are assigned accusative case.



In (27a, b), the embedded IP forms its own case-tier independently of the matrix case-tier since the head INFL of the embedded IP shows subject-verb agreement. In (27a), the syntactic cases of the matrix case-tier are not realized since there are no NPs into which they are mapped. In (27b), the syntactic cases of the embedded IP are mapped into the subject PRO_i and the object *rahan*, respectively, and the nominative case of the matrix case-tier is mapped into PRO_j , which is the only NP in the tier.

5. Conclusion

In this paper, we have shown that case alternations in Finnish infinitives are better accounted for by the feature-based definition of case domain than Maling's (1992) category-based definition. We have proposed that a case

²³. It may also be argued that the nominative case of PRO_i spreads to PRO_j .

²⁴. Note that the verb and PRO raise to INFL and [SPEC, IP], respectively, because there is agreement between them.

domain is licensed by agreement in INFL. The feature-based definition of case domain well accounts for inflected infinitives, which remain problematic in Maling (1992). The feature-based definition need not assume different categories for Finnish infinitives as in Maling (1992) and Vainikka (1989). This paper has focused on infinitives in Finnish. Further studies are needed to deal with case alternations in participle constructions and sentential arguments.

References

- Chomsky, N. (1981) *Lectures on Government and Binding*, New York: Foris.
- Chomsky, N. (1986a) *Barriers*, MIT Press, Cambridge MA.
- Chomsky, N. (1986b) *Knowledge of Language*, Praeger, New York.
- Chomsky, N. (1992) *A Minimalist Program for Linguistic Theory*, MIT Working Papers in Linguistics.
- Lasnik, H. (1992) "Case and Expletives: Notes Toward a Parametric Account", *Linguistic Inquiry* 23:381-405.
- Lee J.-S. (1992) *Case Alternation in Korean: Case Minimality*, Ph.D. dissertation, U Conn, Storrs.
- Maling, J (1992) "Of Nominative and Accusative: the Hierarchical Assignment of Grammatical case in Finnish", to appear in A. Holmber & U. Nikanne (eds.), *Topics in Finnish Syntax*, Mouton de Gruyter, pp. 51-76.
- Pollock, J. Y. (1989) "Verb Movement, Universal Grammar, and the Structure of IP", *Linguistic Inquiry* 20:365-424.
- Timberlake, A (1985) "The Nominative Object in Finnish", *Lingua* 35:201-230
- Vainikka, A (1989) *Defining Syntactic Representation in Finnish*, Ph.D. dissertation, U Mass Amherst.
- Yip, Maling, & Jackendoff (1987) "Case in Tiers", *Language* 63: 217-250.
- Zaenen, Maling, & Thrainsson (1985) "Case and Grammatical Functions: the Icelandic Passive: *Natural Language and Linguistic Theory* 3:441-483.

The English Subjunctive and Minimalist Case Checking Theory*

Kyu-Hong Hwang

1. Introduction

It has been claimed in the minimalist syntax outlined in Chomsky 1993 and 1995b that nominative case feature is checked off by finite tense, [+tense].¹ Given this, such sentences as (1a) and (1b) below directly pose a problem to the theory in that the embedded subject bears nominative case, but no finite tense, i.e., case checker seems to be available, as the base form of verbs suggests.

- (1) a. Mary demanded that he be/*was quiet.
b. Dr. Kim insists that she sing/*sings a song.

The problem of the sort arises in "mandative subjunctive" in English, i.e., the complement clause of words that denote demand, recommendation, proposal, resolution, etc. (Quirk et al. 1985: 156).

This paper seeks to resolve such a paradoxical case by proposing that the subjunctive *that*-clause contains a null modal that carries finite tense. It is also argued that the presence of such an auxiliary prevents overt raising of aspectual verbs like *have* and *be* and the occurrence of the dummy verb *do* in the negative subjunctive.

* I would like to thank Fritz Newmeyer, Heles Contreras, Soowon Kim, Kathy Hunt, Michael Gammon, James Lyle and Yongkil Jeong for their helpful comments.

¹ Chomsky (1995b) believes that this is universal and tries to eliminate AGR-SP on a par with AGR-OP completely. As argued in Hwang 1996a, however, this is very unlikely to be true in that the occurrence of nominative case is sensitive to the presence of AGR-S in a good number of languages that display somewhat marked clause types: untensed but agreed clauses and tensed but unagreed clauses. In this respect, it may be the case that the nominative case checker is parametrized: [+T] in English-type languages and [+AGR-S] in Portuguese-type languages.

The organization of the paper is as follows: section 2 discusses the finiteness of the subjunctive; with the introduction of a non-overt modal, section 3 deals with uninflected verb morphology and peculiar negation phenomena; section 4 recapitulates minimalist case checking theory and shows how the subject of subjunctive clauses has its case feature checked off.

2. Finite Nature of the Subjunctive

It seems clear that the subjunctive exhibits properties of a normal finite clause except for the inflection of the verb. Putting aside the mystery of the verb morphology until next section, let us call our attention to the evidence that suggests that subjunctives are finite clauses.²

A piece of evidence comes from the distribution of [-wh] complementizers: *that* and *for*. As discussed in such work as Chomsky and Lasnik 1977, the former introduces finite clauses, while the latter non-finite clauses (cf. Muysken and van Riemsdijk 1986 for German and Dutch). This fact is borne out by the contrast in grammaticality between a- and b-sentence in (2) and (3) below:

- (2) a. I think that Tom likes you.
 b. *I think for Tom likes you.
- (3) a. For John to study linguistics would surprise me.
 b. *That John to study linguistics would surprise me.

Based on the above fact, it is plausible to say that the subjunctive is a finite clause since it takes *that* rather than *for* as its complementizer. The following example confirms this claim:

² Throughout this paper, the term *finite clause* refers to the clause whose INFL contains finite tense and agreement and alternates with the term *tensed clause*.

- (4) a. It is essential that Mary go alone.
 b. *It is essential for Mary go alone.

Another piece of evidence can be drawn from the parallel behavior of anaphors in subjunctives and finite clauses. It is well-known that anaphors are prevented from occurring in the subject position of finite clauses. This ban has been known as the Tensed S Condition in Chomsky 1973 and the Nominative Island Condition in Chomsky 1980. To illustrate, consider the following examples:

- (5) a. *Students_i believe that themselves_i are good linguists.
 b. *The couple_i expected that each other_i would pass the exam.

Both (5a) and (5b) are filtered out in violation of binding theory condition A outlined in Chomsky 1981 and 1986. According to the condition, anaphors must be bound in what Chomsky (1986: 169) calls "Complete Functional Complex" viz. a minimal governing category that contains a governor, a binder, and a bindee. Given this, the ungrammaticality³ of (5a-b) indicates that the binding domain of anaphors is the embedded clause. This is in fact true in the sense that the subordinate clause has the governor AGR, the bindee *himself* in (5a) and *each other* in (5b), and the binder AGR. Of importance here is Chomsky's (1981, 1986) proposal that AGR in INFL is an accessible SUBJECT to or a binder of anaphors in finite clauses, but not in non-finite clauses. This is borne out by the fact that the governing category is extended to the matrix clauses when the clause in consideration lacks finite INFL, as the well-formedness of (6a-b) shows:

³ Note that Chomsky (1986) later accounts for the ungrammaticality of NIC cases like (5) and (7) in terms of ECP violation under the assumption that anaphors undergo LF movement and adjoin to a higher INFL subject to ECP.

- (6) a. People_i believe themselves_i to be good linguists.
 b. The couple_i expected each other_i to pass the exam.

In short, the coreference across clauses is entirely dependent upon the finiteness of the clause that contains anaphors. To the extent that such an observation is true, we seem to have another criterion that determines the finiteness of subjunctives. The null hypothesis is that if subjunctives are finite, they are expected to behave like finite clauses with respect to the binding phenomena of subject anaphors. With this in mind, now let us look at relevant examples:

- (7) a. *Students_i suggest that themselves_i be good linguists.
 b. *The young girls_i insisted that each other_i take the exam.

As expected, subject anaphors in the subjunctive can not be free within the clause. This amounts to saying that the embedded clause is a governing category for anaphors in (7) (cf. footnote 6). If this is the case, we then can say that the INFL of subjunctives contrasts with non-finite INFL and is exactly the same as that of finite clauses. This argument in turn counts as evidence for the finiteness or tensedness of the subjunctive.

The distribution of sentential adverbs seems to provide additional evidence. As Akmajian (1984) notes, the occurrence of sentential adverbials is limited to tensed clauses. The ungrammaticality of (8a-d) may be due to such a ban. All the sentences below, namely, the imperative (8a), Akmajian's Mad Magazine sentence (8b), the infinitive (8c), and the gerund (8d) are all instances of untensed clauses:

- (8) a. *Fortunately find a good job!
 b. *(What!) Him unfortunately miss the bus!
 c. *For them to definitely leave the room would be awful.

- d. *I was grateful for them certainly knowing the answer.

In contrast to (8a-d), (9a-d), all of which are tensed clauses, allow sentential adverbs. Compare (8) and (9) in light of grammaticality contrast:

- (9) a. I fortunately found a good job.
b. Mary unfortunately missed the bus.
c. That they definitely left the room was awful.
d. I was grateful that they certainly knew the answer.

It is clear from the argument above that there is a correlation between tensedness and the occurrence of sentential adverbials. Such a close bond between the two can in turn serve as one of the criteria that separate tensed clauses from untensed clauses.

Now it is in order to see to how the subjunctive behaves with respect to sentential adverbials. Take a close look at the subjunctive along the line discussed so far:

- (10) a. Mary suggested that we possibly leave the room.
b. I insisted that the sentence definitely be judged grammatical.
c. ?We request that you certainly be quiet in the room.

Most of the native speakers of English I have consulted agreed that (10a-c) are not as good as (9a-d), but better than (8a-d). To the extent that (10a-c) sound awkward at worst, there is a contrast in grammaticality between (8) and (10). The oddness of (10) in contrast to (9) appears to be due to the semantic incompatibility between sentential adverbs and verbs that introduce subjunctive clauses.⁴ Newmeyer (p.c.) also points out that when one makes a

⁴ One may argue that this claim extends to such imperatives as (8a) to the extent that early generative linguists such as Katz and Postal (1964) and Stockwell et al. (1973) treat subjunctives

suggestion, demand, etc., there is no point in expressing adverbial modification in the sentences. If this line of account is on the right track, we can still maintain that subjunctives are finite and license the occurrence of sentential adverbs. The marginal well-formedness of (10a-c) is not because the subjunctive is untensed but because the occurrence of adverbs in question is also subject to semantic and pragmatic conditions, besides syntactic requirements. This kind of conclusion lends additional support to the tensedness of subjunctives.

An argument to the same effect can also be made based on the fact that the subject of the subjunctive is always overt, obligatory and nominative in form. Nominative subjects are always found in tensed clauses in English and at the same time, non-overt subjects occur only in tenseless clauses like those in (13) (cf. Haegeman 1990 and Hwang 1995 for diary contexts). Under some circumstances, however, such null subjects can alternate with overt subjects. When this is the case, the subjects must be accusative in form. These properties are well illustrated below:

- (11) a.⁵ (You and him/*he) never fight again!
 b. What! (Him/*he) get a job!

and imperatives alike. If we take this into account in the present analysis, it may be the case that (8a) violates both syntactic and semantic conditions on sentential adverbials, while (10) fails to meet only the latter condition and hence sounds better than (8a). In fact, it is of interest to note that Katz and Postal (1964) consider imperatives like **Certainly drive the car!* ungrammatical, but I found some speakers who do not completely rule it out. This leaves us room for further research.

⁵ Beukema and Coopmans (1989) assume that the imperative subject is nominative. There is no evidence for their assumption in that we can not tell whether typical imperative subjects such as *everybody, someone, you*, etc. are nominative or accusative. Along the lines suggested in Hwang 1996b, I take the position that the subject of the imperative bears default accusative case based on the accusative form of conjoined subjects and untensedness of the imperative INFL. Refer to Hwang 1996b for more detailed discussion.

- c. (Them/*they) knowing the answer surprised me.
- d. (For him/*he) to leave the room would be great.

As shown in (12a-b), the subject of the subjunctive must be neither null nor accusative in form. Instead, it resembles the subject of the tensed clause in form and occurrence. This fact leads us to treat the subjunctive as a finite clause:

- (12) a. Our decision was that he/*him run for the president.
- b. *Our decision was that [ec] run for the president.
- c. People know that I/*me am innocent.
- d. *People know that [ec] am innocent.

Our discussion for the treatment of the English subjunctive as finite has so far heavily relied on syntactic phenomena somewhat particular to English. The subjunctive clause not being restricted to English alone, one may argue for the same effect on crosslinguistic grounds (Radford 1988). It is reported that subjunctive clauses, like indicative finite clauses, are indeed inflected for tense and agreement in languages with richer morphology than English.⁶ Among such languages is Spanish. Observe the following example taken from Radford (1988: 291):

- (13) a. Exigo [que Juan *parta* para Hawaii Manana]
 I-demand that Juan leaves (3sg.pres. sub.) for Hawaii tomorrow

⁶ It has been claimed in Picallo 1984 and Progovac 1993 that tense in the subjunctive clause is anaphoric to tense in the matrix clause in some languages and thus domain extension to the matrix clause is possible for reflexivization, pronominalization and negative polarity. However, it is important to note that tense in the subjunctive has its own ability to check nominative case of its subject, which is the major concern of our discussion.

b. Exigi[que Juan *partiese* para Hawaii el dia siguiente]

I-demanded that Juan left (3sg.pas.sub.) for Hawaii the following day.

The Spanish subjunctive clause, which is marked by the bracket, is like the finite clause in terms of inflection. The same is true for languages like Rumanian, which shows overt inflection for tense and agreement in the clause under discussion (Grosu and Harvath 1984).

A cautionary word is in order at this point. In a sense, it may be the case that our argument based on Spanish data does not sound so strong as preceding ones to the extent that there is no such a principle that says one particular construction behaves alike in all natural human languages. It must be stressed, however, that crosslinguistic data like (13) don't contradict our claim and instead they tell us how subjunctive clauses would look like if English had richer morphology. In this connection, it is not implausible to consider such crosslinguistic data as supporting evidence for our claim.

In conclusion, it is certain that subjunctive clauses in English are tensed clauses in a number of respects, despite the fact that verbs in the clauses remain uninflected.

3. A Null Modal and its Effects on Verb Inflection and Negation

This section concerns the mystery left unsolved in the preceding section, namely, non-inflecting verb morphology. To proceed, let us pay attention to some English-particular syntactic phenomena that may help us find a solution to the puzzle. One such phenomenon is verb raising. It has been discussed in the literature, for example, Emonds 1978, Pollock 1989, and Chomsky 1991 and 1993 that English differs from French in that the latter triggers overt raising of any main verbs, but not the former.⁷ A good example comes from Pollock 1989:

⁷ A couple of proposals have been advanced as to what forces overt and covert verb raising. Pollock (1989) appeals to the distinction between opaque AGR and transparent AGR in terms of

- (14) a. John often kisses Mary.
 b. *Jean souvent embrasse Marie.
 c. *John kisses often Mary.
 d. Jean embrasse souvent Marie.

The contrast in grammaticality between (14a) and (14b) on the one hand and (14c) and (14d) on the other follows that inflected lexical verbs in French, but not in English, raise overtly over VP-adverbs like *souvent* "often" that are assumed to remain in situ, viz. the VP-adjoined position. (15) shows, however, that overt verb raising takes place even in English when verbs in question are *be* and *have*:

- (15) a. Mary has completely read the book.
 b. *Mary completely has read the book.
 c. I am always ready to help the poor.
 d. *I always am ready to help the poor.

Given the above observation, a descriptive generalization is that ordinary French verbs and English aspectuals *be* and *have* must precede VP-adverbs at least in tensed clauses. Having concluded that English subjunctive clauses are tensed, we expect the generalization to hold in subjunctive clauses. Surprisingly, it turns out that such a prediction is wrong:

- (16) a. The law requires that bills always be paid on time.
 b. *The law requires that bills be always paid on time.

theta-role transmission of verbs to the arguments. On the other hand, Chomsky (1993) relies on the strength of features of INFL. Neither of these two is completely free from criticism on empirical and theoretical grounds, however, as Marantz (1995) points out. What is relevant to the current discussion is a descriptive fact concerning verb raising rather than an explanation for such raising. In this sense, we will stay from what is a driving force.

- c. I insist that he definitely have finished by tomorrow.
- d. *I insist that he have definitely finished by tomorrow

The fact that *be* and *have* do not raise overtly in (16b) and (16d) respectively puts us in a dilemma. This in turn seems to cast some doubt on our previous argument for the tensedness of subjunctive clauses. To make our conclusion and the generalization tenable under given circumstances, we must look for an explanation for the unpredicted behavior of aspectuals in regard to raising in the subjunctive. In fact, there seems to be a way out. Observe the following sentences:

- (17) a. *I will be always ready to help the poor.
- b. I will always be ready to help the poor.

- (18) a. ?*Mary will completely have read the book.
- b. Mary will have completely read the book.
- c. *Mary completely will have read the book.

What we learn from (17) and (18) is that *be* does not behave exactly like *have* in terms of raising over a VP-adverb in sentences that contain a modal auxiliary. The difference is that the former never raises over a VP-adverb in the presence of a modal, while the latter may precede such an adverb in the same environment. At this stage, it is worth noting that *be* and possibly *have* pattern alike in light of raising over a VP-adverb both in subjunctive clauses and in ordinary finite clauses that contain a modal auxiliary.⁸

⁸ It seems that an argument concerning verb raising based on the position of VP-adverbs is not so clear and convincing as it first appears. In particular, such adverbs can enjoy their freedom to some extent in clauses that have a modal and an aspectual *have*, as in (18). Under the present discussion, (18a) rather than (18b) is predicted to be perfectly well-formed, contrary to the fact. This remains unexplained. Nevertheless, I would like to maintain that no aspectuals raise in

This parallel pattern is also observable in relation to the negation *not*. As discussed in the literature cited earlier, only auxiliary verbs and modals, along with negation particle *not*, form negative sentences without employing the dummy verb *do*, as shown in (19). This is attributed to their property to raise overtly to the INFL and principles of economy in the sense of Chomsky 1991. This amounts to saying that non-lexical verbs in English must raise whenever possible, preceding both *not* and adverbs of frequency like *always*. The contrast in the minimal pair of (19) and (20) then receives a natural account:

- (19) a. I am not aware of the fact.
 b. Mary has not finished the work.
 c. *John loves not Mary.
- (20) a. *I do not be aware of the fact.
 b. *Mary does not have finished the work.
 c. John does not love Mary.

As far as French is concerned, the same reasoning forces us to expect that not only auxiliaries but also lexical verbs show the same effect. Indeed, this proves to be the case, as can be seen in (21) and (22) from Pollock 1989 and 1994 respectively:

- (21) a. Jean (n') aime pas Marie.
 Jean (ne) love Neg Marie.
 "Jean does not love Marie."
 b. *Jean ne pas aime Marie.
 Jean (ne) Neg love Marie.

the presence of a modal. For more discussion on adverbs, see Jackendoff 1972, Pollock 1994 and references cited there.

- (22) a. Jean n'a pas disparu.
 "John has not disappeared."
 b. *Jean ne pas a disparu.
 c. Jean n'est pas heureux.
 "John is not happy."
 d. *Jean ne pas est heureux.

It is also the case in English that the occurrence of *do* in the negative is restricted to tensed clauses.⁹ This has been described as Do-support since Chomsky 1957. As an auxiliary, *do* never occurs in constructions like (23), which are deprived of finite tense:

- (23) a. For him (*do) not to go there would disappoint Sue.
 b. I was surprised by John (*does) not knowing the fact.
 c. (What!) Him (*does) not leave early!

With this much background, let us take a close look at negative subjunctive clauses. One of the key features of negated subjunctives is that aspectual verbs never raise over the negative particle *not*, and at the same time, no rescue operations like Do-support take place. This gives us the word order of the negator *not* followed by aspectual verbs, not the other way around. The following examples are to the point:

- (24) a. We suggest that the rules not be abolished.
 b. *We suggest that the rules be not abolished.
 c. *We suggest that the rules do not be abolished.

⁹ Note that *do* also occurs to rescue bound morphemes like an imperative affix IMP, as non-finite negative imperative like *Don't be stupid !* shows (cf. Lasnik 1981).

- (25)¹⁰ a. I demand that Gary not have left by the time it starts.
 b. *?I demand that Gary have not left by the time it starts.
 c. *I demand that Gary does not have left by the time it starts.

The grammaticality shows a direct contrast between sentences in (19) and those in (24) and (25). In other words, what is grammatical in subjunctive clauses is ungrammatical in indicative sentences. With respect to non-occurrence of the expletive *do* in (24) and (25), one may think of two possible accounts. One is due to the complementary distribution between *do* and aspectuals *be* and *have*, and the other is the co-occurrence requirement between *do* and finite tense. Neither of these two is right: the first is ruled out by the fact that in subjunctives, the occurrence of *do* does not depend on the absence of aspectuals, as (26) shows; the second is not worth considering, given our discussion for the tensedness of the clauses in question.

- (26) a. The school requires that students (*do) not play cards.
 b. Tom recommends that we (*do) not go fishing today.

As we have seen in the placement of VP-adverbs (cf. footnote 8), the same situation arises in (24), (25), and (26). The mechanism used in negating subjunctives does not follow the general principles that systematically govern negation operations in other clauses. That is, no verb raises over the negator *not* in subjunctives whether it is lexical or aspectual. In this respect, it is not impossible to consider such a deviant negation operation as exceptional to the general principles. This line of conclusion may be an option, but care must be taken before adopting it in the sense that within framework of generative grammar, it is much more desirable to explain facts than merely to describe

¹⁰ (25a-b) are taken from Johnson (1994:2). The parallel examples are also found in Beukema and Coopmans 1989:

- a. He suggested that he not have finished the work before 10 o'clock.

them. In other words, an exception to the general picture must be considered as a last resort. We are thus forced to take another careful look at the above issue prior to making a final decision.

As a matter of fact, another account can be considered based on the complementary distribution between *do* and a modal:

- (27) a. Mary and John may (*do) not pass the exam.
 b. Both Sue and Jim can (*do) not ride a bicycle.
 c. Those students will (*do) not be surprised.
 d. They will (*do) not have reached the point by next June.

As can be seen from (27c-d), there is an additional fact to which attention needs to be paid. That is, aspectual verbs do not raise over the negator *not* even in negative declarative sentences when they occur with modals. The violation of this results in ungrammaticality of the otherwise well-formed sentences. In short, what seems clear at this point is that modal verbs block Do-support and raising of aspectuals in ordinary negative constructions.

Bearing this observation in mind, now let us compare (24), (25), and (26) with (27). Surprisingly, there is an exact parallelism in verb raising between subjunctives and ordinary clauses with a modal verb. The similarity is so systematic that the parallelism is very unlikely to be a coincidence. Based on this fact, I propose that there exists a phonologically null modal in the subjunctive.¹¹ As long as this claim is viable, all the problems noted so far automatically disappear and receive a unified account. It is of great importance to note that the proposal gains some support in a number of respects. It accounts for otherwise unexplained phenomena unique to subjunctives, for example, peculiar placement of VP-adverbs and the negative

b. *He suggested that he have not finished the work before 10 o'clock.

¹¹ It happens that this proposal is not new. Roberts (1985: 40) makes the same proposal in his footnote 12 along similar lines.

particle *not* and the lack of Do-support. Moreover, it is consistent with our treatment of subjunctives as finite. It appears that the occurrence of modals, just like that of the dummy verb *do*, is limited to tensed clauses. Such a condition on modals is in conformity with the postulation of an empty modal in subjunctive clauses. The starred version of non-finite sentences in (28) shows the restriction on the occurrence of modals

- (28) a. Him (*can) driving the car is well-known to people.
 b. For Tom to (*may) leave would be desirable.
 c. (You all) (*will) go there tomorrow!
 d. What! Me (*will) be a chair!

An empirical fact concerning inversion constitutes further evidence for the above proposal. When negative phrases in the sentence prepose, inversion of an auxiliary verb becomes obligatory. (29) illustrates this point:

- (29) a. *Few people I trusted in the town.
 b. Few people did I trust in the town
 c. *On no account this safe must be opened.
 d. On no account must this safe be opened.

Roberts (1993:323-324) observes that negative preposing induces obligatory auxiliary inversion not only in matrix clauses like (29) but also in subjunctives and in complements of bridge verbs like (30). Consider relevant examples below from Roberts 1993:

- (30) a. ?*I said that in no circumstances he would do that.
 b. I said that in no circumstances would he do that.
 c. ?*I require that under no circumstances he should do that.
 d. I require that under no circumstances should he do that.

- e. I require that under no circumstances he do that.

As in the case of matrix clauses, negative preposing without auxiliary inversion induces ungrammaticality of such sentences as (30a) and (30c). By contrast, the same operation in corresponding sentences (30b) and (30d), where inversion of an auxiliary takes place on a par with preposing of negative elements, gives rise to well-formed sentences. Given this, consideration of (30e) in the same vein makes it inevitable to assume that there exists a null modal such that it acts as a clause operator. Unless we recourse to the existence of an empty modal, the grammaticality of (30e) can not be explained in a principled way. In this regard, the proposal under consideration can be further justified.

A word of caution is now in order. Our discussion so far has been based on the view that the infinitive form of verbs are used in the mandative subjunctives. The fact is more complex than this, however: the modal *should* may occur, as witnessed by (30c-d).¹² This seems to be a dialectal variation. On this matter, Quirk et al. (1985: 1013) state that "In BrE, putative *should* with the infinitive is far more common." If this is indeed the case, the current analysis then becomes more appealing, a welcome result. In the present context, the use of *should* in the subjunctive may be interpreted in such a way that the null modal is phonetically realized in some dialects. That no modal other than *should* appears in the clause may be due to the fact that *should* is the only modal that is compatible with the empty modal in meaning. In case of other dialects like American English, where no overt modal is allowed, there seem to be a straightforward account: the null modal displaces other modals since two modals never occur together in the same sentence.

¹² Roberts(1985) originally does not take (30c-d) into account in his analysis, taking the view that overt modals are impossible in the mandative subjunctive. Quirk et al. (1985: 1014),

We have hitherto appealed to synchronic evidence for positing an empty modal in subjunctives. However, diachronic evidence to the same effect can be made available by looking at subjunctives in the context of the history of the English language. It is said in Turner 1980 and Roberts 1993 that subjunctives distinguished themselves from indicatives by their endings: *-e* in the singular and *-en* in the plural in Old English and Middle English. Such distinctive inflections, along with indicative and infinitive endings, soon started to disappear in Middle English due to phonological changes, so that the subjunctive came to assume an uninflected bare verb by Early Modern English. Roberts (1985, 1993) takes the position that this loss of inflection gave rise to the auxiliary use of modals (cf. Lightfoot 1979 and Denison 1993). That is, when different endings that expressed the mood of the clause became unavailable, modals were reanalyzed as auxiliaries and began to take over the role of inflection, namely, representing modality. Citing Visser's comment, Roberts (1985:41) goes on to say that "It seems that the subjunctive could be replaced by a modal in every major use of the subjunctive." In short, there is a correlation between the erosion of subjunctive endings and the appearance of modals as auxiliaries. Such a reanalysis was inevitable to compensate the loss of inflection and is a natural trend in a view of language change. With all these facts considered, it becomes clear by now that the postulation of an empty modal in subjunctive is neither accidental nor unmotivated even in diachronic perspective.

There is one more aspect that is worthwhile to consider. As Roberts(1993) reports, verbs in the subjunctive still raised to INFL in late Middle English and Early Modern English. This means that such verbs were inflected on the ground that only inflected verbs tend to raise overtly if possible (Chomsky 1993, 1995b and Pollock 1989, 1994). The inflection, however, was deprived of overt morphological realization. Consider the relevant examples (31a-b) cited

however, go on to say that indicative forms are occasionally used in subjunctives like (30c-d) in British English.

in Roberts 1993 and compare them with French subjunctives (31c-d) from Ahn 1994:

- (31) a. And gif he be noght so, then . . .
 And if he be not so, then . . .
 (1420s: James I, Kingis Quair, 62; Gray 1985: 73)
- b. Beware that thou bring not my son thither again.
 (1611: Bible, Gen 24, 6; Visser 1973, § 869)
- c. Je suggere qu'il ne perde pas de temps.
 "I suggest that he not lose any time."
- d. *Je suggere qu'il ne pas perde de temps.
 "I suggest that he lose not any time."

There is one thing in common between (31a-b) and (31c-d), namely, overt verb raising, which are in direct contrast to (23) and (24). The only difference is that morphologically, the inflection is covert in the former, while overt in the latter. One may wonder at this point why verbs raise in (31a-b), given our previous claim that modals came to appear in subjunctives as a new means of expressing modality around late Middle English or Early Modern English. To put it another way, why isn't it the case that a (null) modal is present in (31a-b) and acts as an operator as in (1), blocking verb raising? This may be a reasonable question, but it must be borne in mind that language change is a gradual process, not an overnight phenomenon. It can be inferred from (31a-b) that it was not until Modern English that the null modal completely functioned as an operator in the subjunctive under consideration. This means that there was a transition period between late Middle English and Early Modern English. The fact that (31a-b) had inflection is good enough for the present purpose.

Now let us turn to the issue of uninflected verb morphology. Thanks to the proposal at issue, a natural account is made possible. It is an obvious fact that modals obligatorily carry infection of tense and person agreement,

reflecting only the former in their morphology. This is to say that all the verbs including aspectuals remain uninflected in the presence of modals:

- (32) a. They could do/*did their own work.
 b. Mary may pass/*passes the drive test.
 c. I would be/*am/*was happy if you wash my car.
 d. John should have/*had/*has called the police.

The same scenario extends to subjunctive clauses. Having proposed that there exists a null modal in subjunctives, we expect verbs immediately following the modal to be in base form. By analogy with the behavior of modals in declaratives like (32), we assume that the empty modal carries inflection in subjunctive clauses, viz. tense and agreement, but never manifest it due to its lack of phonetic content. It is for this reason that verbs lack overt inflection in subjunctive clauses like (1), repeated here as (33):

- (33) a. Mary demanded that he be/*was quiet.
 b. Dr. Kim insists that she sing/*sings a song.

To sum up, there are striking similarities between subjunctives and declaratives with a modal in terms of verb raising and verb morphology: aspectuals never raise over VP-adverbs and the negator *not*; verbs including *be* and *have* are in their base form. On the basis of this parallelism, a proposal has been put forward that a non-overt modal exists in the subjunctive, which takes care of all the peculiar syntactic phenomena unique to the clause. We also have justified such a proposal on synchronic and diachronic grounds.

4. Nominative Subjects and Recapitulation of Case Feature Checking

Adopting the basic tenet of Pollock (1989) and developing an earlier idea in his paper (1991), Chomsky (1993) makes a concrete proposal on case checking

of DPs. According to the proposal, nominative case is checked off, via Spec-head agreement, by finite T which head-adjoins to AGR-S. This line of approach underwent a radical revision in Chomsky 1995b, however, where case is reinterpreted as one of the formal features of DPs. With the introduction of new features and subclassification of them into +/- interpretable, Chomsky (1995b) argues that [+interpretable] features are always accessible even after checking, while features with the opposite value are no longer accessible to syntactic processes once checked. The former include categorial- and phi-features of DPs, and the latter case features of DPs, finite T, and transitive V, in addition to phi-and-categorial features of T, V, and A. Chomsky even goes on to say that agreement projections no longer exist at least languages like English and assumes TP in place of AGR-SP (cf. footnote 1).

On nominative case feature checking, Chomsky claims that the nominative case feature is obligatorily and uniformly checked off by finite Tense as a free rider when DPs raises to [Spec TP].¹³ His examples below show how the system works:

- (34) a. seem that John is intelligent
 b. T seems that [TP John T is [AP t₁ [A' t₂ intelligent]]
 c. It seems that John is intelligent.
 d. *John seems that t is intelligent.

Along the lines suggested by Chomsky (1995a, 1995b), we assume that derivation (34a) is formed by two operations: *merge* and *move*. (34b) shows the derivational process of (34a). There are two options in completing derivation (34a): one is merging the derivation with the expletive *it* (34c), and

¹³ What is moving in Chomsky 1995 is not a category but a set of formal features of a lexical item, hence Move F. It is a D-feature of DP that raises to satisfy T's EPP feature, but PF convergence requires the whole category to move along with the feature.

the other is moving *John* to [Spec, TP] of a higher clause (34d). As can be seen from the well-formedness of (34c), only the first option gives rise to a convergent derivation. The reason for this is already given above. When *John* raises to [Spec, TP] of the embedded clause to satisfy the EPP or D-feature of the T, which is [-interpretable] and thus requires obligatory checking, its phi- and nominative case-features, along with its D-feature, are checked by the same T as a free-rider. When *John* further moves to [Spec, TP] of the matrix clause, it checks off D-and phi-features of the matrix T, but not T's case assigning feature. This is because unlike its other [+interpretable] features, the [-interpretable] case feature of *John* at this point is erased as result of its checking with the lower T and does not enter into further checking relation with the higher T. That is, the nominative case assigning feature of the T, which must be checked, being [-interpretable], is left unchecked and consequently constitutes an illegitimate object at LF. This explains why derivation (34d) crashes. On the other hand, when we merge structure (34a) with the expletive *it*, all the features of the matrix T are checked off. Thus, no illegitimate object exists at LF and convergent derivation (34c) emerges.

The exactly opposite situation arises when the subordinate clause is non-finite, as shown in (35):

- (35) a. seems John to be intelligent
 b. John seems t to be intelligent.
 c. *It seems that John to be intelligent.

At the point where derivation reaches (35a) in the way depicted in (34b), raising *John* to the higher [Spec, TP] leads to convergent derivation (35b). This follows that *John* can not have its nominative case feature checked off in the lower [Spec, TP] since the T is non-finite and hence lacks nominative case-assigning feature. Given that the [Spec, TP] of the main clause is the only place where its case feature can be checked off, *John* has no choice but to move up to the position. In the same configuration, checking of the rest of

features are done at the same time. The ungrammaticality of (35c) is straightforward: *John's* case feature remains unchecked. Moreover, it is out of the question to check the case feature at LF. T's case feature in the matrix clause is no longer available since it already entered into checking relation with the expletive *it* and disappeared due to its [-interpretable] nature.

Now turning to case checking of the subjunctive subject¹⁴ in (1), reproduced here as (36), we assume that checking proceeds along the lines outlined above.

- (36) a. Mary demanded that he be/*was quiet.
 b. Dr. Kim insists that she sing/*sings a song.

A close look at the derivational process of the subjunctive gives a clearer picture. The process is given in (37):

- (37) a. Mary demanded that he T- \emptyset be [_{AP} t₁ [_{A'} t₂ quiet]]
 b. Dr. Kim insists that she T- \emptyset [_{VP} t sing a song]

As in the case of (34) and (35), the strong categorial feature of T triggers overt movement of the subject, *he* in (37a) and *she* in (37b) to [Spec, TP]. The raised DP then enters into checking relation with the host, T and all the relevant features such as phi-, case-, and categorial-features of both sides are checked off at once, hence a convergent derivation. It must be borne in mind that the analysis heavily relies upon the previous proposal that T, carried by an empty modal \emptyset , in the subjunctive is finite and has a nominative case-checking ability. Without such a proposal, the well-formedness of (36) would be left unexplained in a case-theoretic point of view. We therefore must not

¹⁴ Beukema and Coopmans (1989) claim that the imperative and the subjunctive in English has an INFL of [-Tense, +AGR] and that [+AGR] assigns nominative case to the subject of both

be misled by the uninflected verb morphology and arrive at the wrong conclusion the subjunctive has non-finite tense.

In sum, we have recapitulated the Case Feature Checking Theory proposed in Chomsky 1995b with an emphasis on nominative case. It has also been shown that the subject of subjunctives, just like that of indicatives, has its nominative case feature checked off against the finite T in the clause when the strong categorial feature of the T attracts the subject overtly.

5. Conclusion

In an attempt to account for the case checking of the subject in the mandative subjunctive in the minimalist case theory, a claim has been made in this paper that subjunctives are finite although the uninflected verb morphology suggests otherwise. Section 2 has provided several pieces of evidence for the claim: the subjunctive is introduced by the complementizer *that* rather than *for*; the clause displays NIC effects and may co-occur with sentential adverbs; the subject in the sentence must be overt and nominative in form; the subjunctive assumes the same inflection as the indicative in languages with rich morphology. In order to take care of the uninflected verb morphology and the absence of both Do-support and aspectual verb raising, section 3 has advanced a proposal that there exists an empty modal in subjunctives. It has also been demonstrated that such a proposal is synchronically and diachronically well motivated. The last section has laid out the Minimalist Case Feature Checking Theory and illustrated how the subjunctive subject gets its case feature discharged in the theory.

clauses. Yet, this is at odds with the current minimalist approach (cf. Hwang 1996b for the imperative).

References

- Ahn, Hee-Don. 1994. Remarks on head movement in English and French. In Explorations in generative grammar, ed. The preparatory Committee for "A Festschrift for Dong-Wee Yang", 1-25. Seoul: Hankuk Publishing Co.
- Akmajian, Adrian. 1984. Sentence types and the form-function fit. Natural Language and Linguistic Theory 2: 1-23.
- Beukema, Frits and Peter Coopmans. 1989. A government-binding perspective on the imperative in English. Journal of Linguistics 25: 417-436.
- Chomsky, Noam. 1957. Syntactic structure. The Hague: Mouton.
- Chomsky, Noam. 1973. Conditions on transformations. In A festschrift for Morris Halle, ed. Stephen R. Anderson and Paul Kiparsky, 232-286. New York: Holt, Rinehart and Winston.
- Chomsky, Noam. 1980. On binding. Linguistic Inquiry 11: 1-46.
- Chomsky, Noam. 1981. Lectures on government and binding. Dordrecht: Foris.
- Chomsky, Noam. 1986. Knowledge of language: its nature, origin, and use. New York: Praeger.
- Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Principles and parameters in comparative grammar, ed. Robert Freidin, 417-457. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger, ed. Kenneth Hale and Samuel Jay Keyser, 1-52. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1995a. Bare phrase structure. In Government and binding theory and the minimalist program, ed. Gert Webelhuth, 383-420. Cambridge, Mass.: Basil Blackwell.

- Chomsky, Noam. 1995b. Categories and transformations. In The Minimalist Program, 219-394. Cambridge, Mass.: MIT Press.
- Chomsky, Noam and Howard Lasnik. 1977. Filters and control. Linguistic Inquiry 8: 425-504.
- Denison, David. 1993. English historical syntax: verbal constructions. New York: Longman Publishing.
- Emonds, Joseph. 1978. The complex V_V' in French. Linguistic Inquiry 9: 151-175.
- Grosu, Alexander and Julia Horvath. 1984. The GB theory and raising in Rumanian. Linguistic Inquiry 15: 348-353.
- Haegeman, Liliane. 1990. Non-overt subjects in diary contexts. In Grammar in progress, ed. Joan Mascaró and Marina Nespor, 167-179. Dordrecht: Foris.
- Hwang, Kyu-Hong. 1995. An empty category in the so-called English bare imperative. Ms., University of Washington.
- Hwang, Kyu-Hong. 1996a. Parametrization of nominative case checking in minimalist syntax. Ms., University of Washington.
- Hwang, Kyu-Hong. 1996b. An overt subject in the English imperative and default case feature checking theory. Ms., University of Washington.
- Jackendoff, Ray. 1972. Semantic interpretation in generative grammar. Cambridge, Mass.: MIT Press.
- Johnson, Kyle. 1994. Head movement, word order and inflection. In Head movement: the interface between morphology and syntax, ed. S.-H. Park and J.-Y. Yoon. Seoul: Hankuk Publishing Co.
- Katz, Jerrold and Paul M. Postal. 1964. An integrated theory of linguistic descriptions. Cambridge, Mass.: MIT Press.
- Lasnik, Howard. 1981. Restructuring the theory of transformations: a case study. In Explanation in linguistics, ed. Norbert Hornstein and David Lightfoot, 152-173. New York: Longman.
- Lightfoot, David. 1979. Principles of diachronic syntax. Cambridge: Cambridge University Press.

- Marantz, Alec. 1995. The minimalist program. In Government and binding theory and the minimalist program, ed. Gert Webelhuth, 349-381. Cambridge, Mass.: Basil Blackwell.
- Muysken, Pieter and Henk van Riemsdijk. 1986. Projecting features and featuring projections. In Features and projections, ed. Pieter Muysken and Henk van Riemsdijk, 1-30. Dordrecht: Foris.
- Picallo, M. Carme. 1984. The INFL node and the null subject parameter. Linguistic Inquiry 15: 75-102.
- Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. Linguistic Inquiry 20: 365-424.
- Pollock, Jean-Yves. Notes on the clause structure. In Head movement: the interface between morphology and syntax, ed. S.-H. Park and J.-Y. Yoon. Seoul: Hankuk Publishing Co.
- Progovac, Ljiljana. 1993. Subjunctive: the (mis)behavior of anaphora and negative polarity. Linguistic Review 10: 37-59.
- Quirk et al. 1985. A comprehensive grammar of the English language. New York: Longman Inc.
- Radford, Andrew. 1988. Transformational grammar. Cambridge: The Press Syndicate of the University of Cambridge.
- Roberts, Ian. 1985. Agreement parameters and the development of English auxiliaries. Natural Language and Linguistic Theory 3: 21-58.
- Roberts, Ian. 1993. Verbs and diachronic syntax: a comparative history of English and French. Dordrecht: Kluwer Academic Publishers.
- Stockwell et al. 1973. The major syntactic structures of English. New York: Holt, Rinehart and Winston, Inc.
- Turner, John. 1980. The marked subjunctive in contemporary English. Studia Neophilologica 52: 271-277.

**The evolution of pronominal ille from Latin to Spanish:
a grammaticalization analysis.**

by José Carrasquel

0. INTRODUCTION This study purports to show that the emergence and evolution of the third-person pronouns in Spanish can be explained as a phenomenon of grammaticalization. Section one outlines the grammaticalization framework and research program. Section two adduces evidence from medieval texts to support the claim that grammaticalization is the culprit for the emergence of the third-person pronouns in Old Spanish. Section three gathers evidence from colloquial Spanish to predict that the evolution of the third-person pronouns into Future Spanish will be the result of further grammaticalization. Section four proposes grammaticalization chains to illustrate the evolution in question, and finally section five offers the conclusions of this study.

Abbreviations

Per language

CL = Classical Latin, FSp = Future Spanish, MFr = Modern French, MRom = Modern Romance, MSp = Standard Modern Spanish, ORom = Old Romance, OSp = Old Spanish, VL = Vulgar Latin

Per bibliographical item

D&D = Díaz y Díaz (1950), Ford = Ford (1934), GS = Glosas Silenses in MP, G&H = Gifford and Hodcroft (1966), HCH = Heine et al (1991), H&R = Heine & Reh (1984), MP = Menéndez Pidal(1956), OD = Other Documents in MP

1. THE RESEARCH PROGRAM This paper deals with the phenomenon in the evolution from Latin to Spanish whereby the pronominal paradigm of the CL remote demonstrative *ille* 'that' in (1) gave way to the five functionally different paradigms in MSp shown in (2).

(1)

	M	F	N
sg.	Nom. <i>ille</i>	<i>illa</i>	<i>illud</i>
	Gen. <i>illīus</i>	<i>illīus</i>	<i>illīus</i>
	Dat. <i>illī</i>	<i>illī</i>	<i>illī</i>
	Acc. <i>illum</i>	<i>illam</i>	<i>illud</i>
	Abl. <i>illō</i>	<i>illā</i>	<i>illō</i>
pl.	Nom. <i>illī</i>	<i>illae</i>	<i>illa</i>
	Gen. <i>illōrum</i>	<i>illārum</i>	<i>illōrum</i>
	Dat. <i>illīs</i>	<i>illīs</i>	<i>illīs</i>
	Acc. <i>illōs</i>	<i>illās</i>	<i>illa</i>
	Abl. <i>illīs</i>	<i>illīs</i>	<i>illīs</i>

(2)

	3rd Person Pronouns Subject/ DO IO Prep.			New Demonstr.	Antecedent Pronouns
	'he/she/ they' (1)	'it/him/her them' (2) (3)		'that one/ those ones' (4)	'the one(s) [who(m)]' (5)
m.s.	<i>él</i>	<i>lo</i> <i>le</i>		<i>aqué</i>	<i>el (que, cual)</i>
f.s.	<i>ella</i>	<i>la</i> <i>le</i>		<i>aquélla</i>	<i>la (")</i>
m.p.	<i>ellos</i>	<i>los</i> <i>les</i>		<i>aquéllos</i>	<i>los (")</i>
f.p.	<i>ellas</i>	<i>las</i> <i>les</i>		<i>aquéllas</i>	<i>las (")</i>
n.s.	<i>ello</i>	- -		<i>aquello</i>	<i>lo (")</i>

The main claim of this paper is that the evolution from the pronominal forms in (1) down to the paradigms in (2) is an example of grammaticalization in the emergence of ORom analogous to other non-Romance phenomena described in Heine and Reh (1984) [H&R] and Heine et al (1991) [HCH]. The key concept in their framework is the notion of grammaticalization, the linguistic mechanism whereby, in the course of time, lexical items acquire new related meanings or grammatical functions which come to convey relations that either were not expressed before or were expressed differently. The assumption underlying this framework is that "the basis of grammaticalization is to be sought outside language structure" [HCH(22-23)], and that grammaticalization is motivated by extralinguistic factors, above all cognition.

An example of grammaticalization in the evolution from CL to MSp not related to this study is given in (3) to illustrate the phenomenon.

(3)

- | | | | |
|----|-----|------------------|--------------------------|
| 1. | CL | domus habeo | 'I have a house' |
| 2. | VL | nihil dare habeo | 'I have nothing to give' |
| 3. | OSp | dare he | 'I have to give' |
| 4. | MSp | daré | 'I will give' |

lexical verb	->	modal verb	->	temporal desinence
POSSESSION		OBLIGATION		FUTURITY

In (3) the lexical verb habere denoting possession in CL as in (3.1) becomes an auxiliary verb denoting obligation in OSp as in (3.3), and turns into a temporal desinence marking future tense in MSp as in (3.4). The process in (3) shows 1) how novel means for marking obligation and futurity arise from the frequent and colloquial usage of the verb habere, 2) how the

lexical item habere turns into the grammatical desinence -é, and 3) how the notions of possession, obligation and futurity become linked by cognition in the course of time.

H&R propose that grammaticalization phenomena result from the interaction of the three kinds of processes in (4).

(4)

1. phonological processes which modify the phonetic substance of a linguistic unit such as adaptation, fusion, erosion and loss.
2. morphosyntactic processes which modify the morphological and/or syntactic status of a unit such as compounding, cliticization and affixation.
3. functional processes which modify either the meaning of a lexical unit or the function of a grammatical unit such as desemanticization, paradigmatic and syntagmatic analogy, semantic split and semantic shift.

The parameters that H&R use to determine whether a given unit or paradigm has undergone grammaticalization are given in (5).

(5)

The more grammaticalization a linguistic unit undergoes:

1. the less semantic complexity, functional significance and expressive value it has;
2. the less pragmatic and the more syntactic significance it has;
3. the fewer members in the same paradigm there are;
4. the less free and the more fixed its position in the sentence is;
5. the more obligatory and general its use becomes in certain contexts;
6. the more it coalesces semantically, morphosyntactically, and phonetically with other units (i.e. the more homonymy)
7. the less phonetic substance it has.

In general, grammaticalization is taking place whenever an old, more concrete, lexical or less grammatical unit develops into a novel, more abstract, more grammatical category. Thus, the hypothesis underlying this investigation is that if ille indeed underwent grammaticalization, then its development from CL to MSp will be shown 1) to be a linguistic continuum which involves the phonological, morphosyntactic, and functional processes in (4), and 2) to comply with the parameters in (5).

2. EMERGENCE OF THE THIRD- PERSON PRONOUNS IN OSp The development of paradigm (1) into (2) can be segmented into three phases: from CL to VL, from VL to OSp, and from OSp to MSp.

2.1. FROM CL TO VL The CL forms in (1) are stressed, polysyllabic pronouns morphologically marked for case, number and gender, and their focal sense is remoteness, spatial, temporal or mental. In the evolution from CL to VL processes of paradigmatic analogy, semantic shift, and syntagmatic analogy can be posited as follows.

2.1.1. PARADIGMATIC ANALOGY Based on the VL excerpts in D&D, the CL pronominal system in (1) must have been reduced to the VL system in (6). The target system exhibits but a remnant of the six case oppositions in the source system, namely the non-dative:dative opposition shown in (6).

(6)

number		sg.	pl.
case			
Non-Dat.	}	M,N	illo illos
		F	illa illas
Dat.		M,F	illi illis

2.1.2. SEMANTIC SHIFT The forms in (6) have undergone a shift in focal sense as follows: the function as a definite marker, which started out as a non-focal sense in early VL is foregrounded and becomes the most prominent use in late VL, while the function as a distal marker, which used to be the focal sense, is backgrounded. The sentences in (7) and (8) below contain units from (6) which are devoid of any semantic nuances of remoteness and solely indicate definite reference.

2.1.3. SYNTAGMATIC ANALOGY The VL derivatives in (6) can be found in subject, indirect object, and mainly in direct object position having ousted the CL determinative *is* 'it:the:that' from these functions as shown by the sentences in (7).

(7) model

VL sentence

M_{Sp} counterpart

English gloss

source year:page#:line# or sentence#

As a subject pronoun:

1. *jn*Clunia fuerunt illos or¹ illos eas-mandarunt
 ellos fueron a Cluña or ellos las mandaron
 'they went to Cluña' or 'they sent them'
 OD 1030:35:43

As a direct object pronoun:

2. ubi illo potueritis inuenire
donde lo pudierais conseguir / donde pudierais conseguirlo
'where you could find it'
OD 1061:26:5
3. partiberunt se illa jillos infancias
se la repartieron los infantes
'the princes distributed it among themselves'
OD 1030:35:16

As an indirect object pronoun:

4. si filios abueris, serbiat illis
si tuvieras hijos, sírveles
if you had children, serve to them
OD 1078:27:17

The pronominal forms in (6) are also found as prepositional pronouns as in (8.1), as antecedents of relative syntagms as in (8.2), and as antecedents of prepositional syntagms as in (8.3) and (8.4).

(8)

1. As a prepositional pronoun:
... exeptis illos
... a excepción de ellos
'... except for them'
OD 1078:27:9
2. As an antecedent of a relative syntagm:
et ilo que fuit de meo pater
y lo que fue de mi padre
'and that which was of mi father'
OD 1097:31:6

As an antecedent of a prepositional syntagm:

3. et illas de auola tua
y las de tu abuela
'and those of your grandmother'
OD 1078:27:18
4. et illo de Dobres
y el de Dobres
'and that from Dobres'
OD 1055:30:15

2.1.4. EVOLUTION UP TO VL The evolution of the pronominal demonstrative forms from CL to VL is given in (9) in terms of a feature diagram. The features² are the properties that the demonstrative unit ille exhibits in CL, and the feature diagram is but an empirical tool useful insofar as it illustrates the gradual and systematic nature of the phenomenon.

(9)

feature stage	dei	def	cas	ton	dis
CL	+	+	+	+	+
VL	+	+	+	+	-

(9) shows that there was a semantic shift of the pronominal forms from marking distance and definiteness in CL to solely marking definiteness in VL.

2.2. FROM VL TO OSp This phase is the result of processes of adaptation and semantic split.

2.2.1. ADAPTATION The OSp derivatives of the pronominal system in (6) are given in (10).

(10)

		Sg.	Pl.	
Non-Dat.	{	M,N	el:elo:lo	elos:los
		F	ela:la	elas:las
Dat.	M,F	li	lis	ge (preceding a d.o.p.)

The system in (10) is composed of tonic and atonic allomorphs which are in complementary distribution as follows: 1) the atonic forms lo, la, los, las, li, lis appear within the verbal syntagm as grammatical objects or as antecedents in relative syntagms as shown in (11)

(11)

As a direct object:

1. dola (ell aldea)
 la doy (la aldea)
 'I give it' (the village)
 G&H 11:30:3

As an indirect object:

2. saquent li los oculos (al latronem)
 sáquenle los ojos (al ladrón)
 'take the eyes out of him' (the thief's)
 G&H 7:26:23

As an antecedent pronoun in an oblique relative syntagm:

3. delos que auien ganados
 de los que hayan ganado
 'from the ones (that) they have won'
 Ford IV:14:101

The tonic forms in (10) are used as subject and prepositional pronouns, and as antecedent pronouns in non-oblique relative syntagms as shown in (12).

(12) model

OSp sentence
 MSp counterpart
 English gloss
 source:page#:line#

1. As a subject pronoun:
 qu'ela la-fiziera
 que ella la hiciera
 'that she did it'
 G&H 18:48:125
2. As a prepositional pronoun:
 et mager quiera ningun de elos
 y quizá no quiera ninguno de ellos
 'and perhaps none of them wants to'
 G&H 13:32:22
3. As a non-oblique antecedent pronoun:
 elas quiet nominata por baptizare
 las que nominadas por bautizar
 those who nominated to be baptized
 GS VIII:18:204

2.2.2. SEMANTIC SPLIT The desemanticization of ille that led to the foregrounding of the definite focal sense and the backgrounding of the original distal focal sense ends in OSp with the functional split of the remoteness and definiteness senses as follows :

1) the original focal sense of remoteness is maintained in OSp by reinforcing³ the non-dative tonic ille derivatives in (10) with the VL preposition *accu 'at here is', which gave rise to new demonstrative pronominal forms in (13).

(13)

1. akelos qui tornaren
 aquéllos que vuelvan
 'those who come back'
 GS XI:21:300

2. Aquelas non las puede leuar
 Aquéllas no las puede llevar
 'Those s/he can not carry them'
 Ford IV:15:116

2) The new focal sense of definiteness becomes the only sense of the OSp derivatives in (10).

2.2.3. EVOLUTION UP TO OSp The feature diagram in (14) shows the changes that took place in the transition from CL to OSp.

(14)

feature	dei	def	cas	ton	dis
stage					
CL	+	+	+	+	+
VL	+	+	+	+	-
OSp	+	+	+	+/-	-

(14) shows that in the evolution from CL to OSp an allomorphy between tonic and atonic forms emerged.

2.3. FROM OSp to MSp This section is divided into the evolution of the disjunctive forms and the evolution of the conjunctive forms in (10).

2.3.1. EVOLUTION OF THE OSp DISJUNCTIVE FORMS The OSp tonic forms in (10) gave rise to the MSp system in (15).

(15)		sg.		pl.
	M	él:el (que)		ellos:los (que)
	N	ello:lo (que)	-	
	F	ella:la (que)		ellas:las (que)

The MSp tonic forms in (15) are used as subject and prepositional pronouns while the atonic forms are used as antecedent pronouns in prepositional and relative syntagms. The atonic forms are the result of a process of apheresis followed by cliticization which changed the OSp antecedent pronouns elo (que), ela (que), elos (que), and elos (que) into the MSp lo (que), el (que), la (que), los (que) and las (que).⁴

2.3.2. EVOLUTION OF THE OSp CONJUNCTIVE FORMS The OSp atonic forms in (10) gave rise to the MSp atonic system in (16).

(16)		sg.		pl.
Non-Dat.	{	M,N	lo	los
		F	la	las
Dat.		M,F	le	les se (preceding a d.o.p.)

The OSp:MSp pairs of sentences in (17) show two interesting facts about the grammaticalization status of the forms in (16) with respect to their OSp atonic predecessors in (10).

(17)

1. cuando las non queredes (mis hijas)
cuando no las querais (a mis hijas)
'when you do not want them' (my daughters)
G&H 14:35:106
2. las han dexadas (sus hijas)
las han dejado (sus hijas)
'they have left them (his daughters)
G&H 14:36:21
3. ouieron los de arrancar
los tuvieron que arrancar / tuvieron que arrancarlos
'they had to take them apart'
G&H 14:34:53

First of all, the fact that the OSp atonic forms can be placed in several positions with respect to the finite verb in the sentences in (17) suggests that even though the OSp forms are already somewhat syntagmatically bound, they have not undergone cliticization. However, their MSp counterparts in (17) seem to have less syntagmatic freedom since they are always adjacent to the verbal syntagm either preceding the finite forms or combined with an infinitive as in (17.3) in declarative sentences.

When the MSp pronoun precedes the verbal syntagm, it is conspicuous that it has undergone cliticization because no other particle can come in between the finite verbal form and the atonic pronoun, and when the pronoun is attached to an infinitive or to a gerund, they form one unit. However, in OSp the atonic pronoun can either be separated from the finite form by the negation particle non as in (17.1), 2) or precede the finite form as in (17.2) or follow the finite form as in (17.3).

Furthermore, in the OSp sentence in (17.2) there are two markers of number and gender in the verbal syntagm, namely the direct object pronoun and the past participle, whereas in its MSp counterpart the only

marker of number and gender in the verbal syntagm is the direct object pronoun. Both of these facts together suggest that the MSp object pronoun is more grammaticalized than its OSp ancestor:

2.3.3. EVOLUTION UP TO MSp The evolution from CL to MSp is shown as a feature diagram in (18) taking only into consideration the conjunctive pronouns in (16).

(18)

feature	dei	def	cas	ton	dis
stage					
CL	+	+	+	+	+
VL	+	+	+	+	-
OSp	+	+	+	+/-	-
MSp	+	+	+	-	-

(18) shows that the object pronouns have become clitics in MSp.

3. EVOLUTION INTO FSp This section deals with three phenomena currently underway in some dialects of MSp which involve the clitic pronouns in (16) and are considered characteristics of FSp.

3.1. LAISMO AND LEISMO Sentence (19.1) illustrates 'laísmo' as the use of the direct object pronoun la in place of the indirect object le, and sentence (19.2) illustrates 'leísmo' as the use of the indirect object pronoun le in place of the direct object pronoun lo with human referents in colloquial Peninsular Spanish, which will be labeled FSp.

(19)

1. FSp No la voy a dar nada
 MSp No le voy a dar nada
 'I won't give her anything'
2. FSp No le he visto
 MSp No lo he visto
 'I haven't seen him'

The establishment of 'laísmo' and 'leísmo' taken together would change the MSp system in (16) as shown in (20).

(20)

	sg.	pl.	
M	}	le	les (human)
		lo	los (non-human)
F		la	las

(20) shows the eradication of the last vestige of the case distinction accusative:dative from the MSp atonic third-person pronominal system in (16). The feature diagram in (18) would change as shown in (21).

(21)

feature stage	dei	def	cas	ton	dis
CL	+	+	+	+	+
VL	+	+	+	+	-
OSp	+	+	+	+/-	-
MSp	+	+	+	-	-
FSp	+	+	-	-	-

(21) shows that the clitic pronouns become unmarked for case in FSp. They become less morphologically marked and more generalized in accordance with grammaticalization.

3.2. INDIRECT OBJECT DOUBLING (22) shows sentences that involve indirect objects.

(22)

1. a. ?Yo _ la di a Mario
b. Yo se la di a Mario
'I gave (him) it to Mario'
2. a. *¿A quién _ diste la flor?
b. ¿A quién le diste la flor?
'Who did you give (him/her) the flower to?'

In all dialects of MSp, a declarative sentence containing an indirect object also contains the corresponding indirect object pronoun; otherwise, the sentence is anomalous such as (22.1a). Furthermore, an interrogative sentence which contains an indirect object and does not have the corresponding indirect object pronoun is completely ungrammatical such as (22.2a). The prominent use of le and les in sentences containing indirect objects shows that these pronouns are being generalized as grammatical markers, more specifically as agreement markers. The sentence in (23) shows even more grammaticalization of the indirect object pronoun le.

(23)

- Le traje dulces a los niños
'I brought (them) candy for the children'

The fact that the sentence in (23) is grammatical in some dialects of MSp even though the indirect object pronoun le has the plural referent los niños shows a paradigmatic reduction from leles to simply le. The le in (23) is no longer a pronominal marker of definite reference as its MSp predecessor, but rather it is a grammatical marker which indicates agreement between the verb and its indirect object, within the verbal syntagm.

3.3. PRONOUN ELISION The phrases in (24) taken from Lope Blanch (1990) exemplify the usage of the clitic pronouns in the U.S. Southwest Spanish dialect.

(24)

- | | | |
|----------------------|-----------------|---|
| 1. sopapía sí l'hago | [sopapiasilayo] | 'sopapilla I do make <u>it</u> ' |
| 2. no l'iba a gustar | [noliβayustar] | 'you were not going to like <u>it</u> ' |

The phrases in (24) show further apocope of the singular clitic pronouns when the verb they cliticize onto starts with a vocalic segment. The elided forms in (24) can be considered allomorphs of the clitics in (20) in FSp.

3.4. EVOLUTION UP TO FSp The case neutralization and the indirect object doubling phenomena together would change the feature specification for FSp in (21) as shown in (25).

(25)

feature	dei	def	cas	ton	dis
stage					
CL	+	+	+	+	+
VL	+	+	+	+	-
OSp	+	+	+	+/-	-
MSp	+	+	+	-	-
FSp	+	-	-	-	-

(25) shows that the FSp clitic pronoun is unmarked for case and devoid of definite reference thus becoming a purely grammatical unit. The clitic pronoun in FSp is still deictic because it is an agreement marker between the verbal unit and its object.

4. A GRAMMATICALIZATION ANALYSIS This section sums up the processes and evidence adduced so far by means of grammaticalization chains. In these chains, CL ille and its MSp derivatives are represented as

units along the same continuum with a few in-between stages and a hypothetical future stage. Given any two units along a chain, it is always the case that the left unit is more lexical than the right one, and the right unit is more grammatical than the left one.

4.1. GRAMMATICALIZATION CHAINS (26) contains four grammaticalization chains⁵, each of which shows the most salient predecessors of the MSp paradigms in (2). In this section, the reinforced derivatives aqué and its forms have been left out.

(26)	CL		VL		OSp		MSp		FSp	
1.	ille	}	illo	->	{	el (que)	->	él, el (que)		
	illīfus					elo (que)	->	ello, lo (que)		
	illum					lo	->	lo	->	l'
	illō									
	illī	->	illi	->	li	->	le	->	l'	
2.	illa	}	illa	->	{	ela (que)	->	ella, la que		
	illīfus					la	->	la	->	l'
	illam									
	illā									
	illī	->	illi	->	li	->	le	->	l'	
3.	illī	}	illos	->	{	elos (que)	->	ellos, los que		
	illōrum					los	->	los		
	illīs									
	illōs									
	illīs	->	illis	->	lis	->	les	->	le	

4.	illae	}	->	illas	->	{	elas (que)	->	ellas, las que
	illārum						las		las
	illīs								
	illas								
	illīs	->	illis	->	lis	->	les	->	le

4.2. PARADIGMATIC CHANGES

This section summarizes the

main paradigmatic changes in (26) which result from the linguistic processes in (4). The validity of this analysis is founded on the argument that every single change exhibited by the chains adheres to the grammaticalization parameters in (5).

In the phase from CL to VL, grammaticalization narrows down morphological case to a twofold distinction of dative:non-dative; however, all the forms maintain their syntagmatic and prosodic autonomy.

Concomitantly, as a product of desemanticization the focal sense of remoteness of the CL paradigm becomes rivaled by a new focal sense of definiteness.

In the phase from VL to OSp, the non-dative forms get further grammaticalized by a process of adaptation resulting in an allomorphic system consisting of tonic and atonic allomorphs. The atonic forms function as direct object pronouns, indirect object pronouns while the tonic forms function as subject and prepositional pronouns, and as antecedents of relative and prepositional syntagms,. Also, a process of semantic split and shift takes place whereby the focal sense of the VL forms becomes definiteness, and new forms maintaining the former functional status as demonstratives are created by reinforcement with *accu.

The tonic OSp forms further split into tonic and atonic forms in their evolution to MSp as follows: the MSp atonic derivatives function as antecedents of relative and prepositional syntagms while the tonic forms are used as subject and prepositional pronouns. The OSp atonic forms lo, la, los, las, li, lis become more grammaticalized as they evolve into MSp by a process of cliticization onto the verbal syntagm.

Interestingly, the indirect object doubling phenomena⁶ suggest that the MSp dative form le will be further grammaticalized as an agreement marker within the verbal syntagm in FSp. Also, 'laísmo' and 'leísmo' phenomena suggest that the case opposition accusative:dative in the MSp atonic pronominal system will disappear in FSp. Both of these predictions for FSp are in agreement with the grammaticalization parameters in (5).

4.3. COGNITIVE LINKS (27) shows the associations in the grammaticalization chains in (26) at the cognitive level. All the units along the chains are deictic insofar as they stand in place of or point to a referent which is a spatial entity, a previously mentioned element in discourse or text, or a grammatical object.

(27)

CL	VL	OSp	MSp	FSp
demonstrative --> pronoun		third-person pronoun	-->	agreement marker

(27) shows that a demonstrative pronoun is recruited as the source unit to express definiteness given the capability of demonstratives to express both remoteness and definiteness simultaneously. In the evolution to OSp the

capacity to express definiteness becomes the focal sense of the pronominal forms (although their capacity to signal remoteness is maintained with reinforced forms). Once the pronominal forms cease to mark distance, they are better defined as third-person pronouns.

The evolution into FSp could continue by recruiting the most grammaticalized pronouns in MSp, namely indirect object clitics, to express agreement between the verb and its indirect object. Once the clitic becomes obligatory in the verbal syntagm and ceases to have definite reference, it can be better defined as an agreement marker.

The diagram in (28) shows the paradigmatic and cognitive associations in this process of grammaticalization.

(28)

CL	VL	OSp	MSp	FSp
demonstrative --> pronoun		third-person pronoun	-->	agreement marker
SPATIAL REFERENCE		DISCOURSE REFERENCE		GRAMMATICAL REFERENCE

(28) shows that the process whereby a demonstrative pronoun turns into a third-person pronoun, a more grammatical unit, is driven by the cognitive metaphor SPACE-TO-DISCOURSE which draws concepts and units from the domain of spatial deixis to express relations in the domain of discourse deixis.

(28) also shows that the process whereby a third-person pronoun becomes an agreement marker, a purely grammatical unit, is driven by the cognitive metaphor DISCOURSE-TO-GRAMMAR which draws concepts and units from the domain of discourse deixis to express relations in the domain of grammatical deixis or grammatical agreement.

5. CONCLUSION Previous accounts point to 'weakening' or 'bleaching' as the process responsible for the evolution from CL ille to the MSp third-person pronouns, thus equating language change with language decay. This study replaces the vague notion of 'weakening' with the principled and systematic notion of grammaticalization, a mechanism motivated by cognition whereby the phonological, morphosyntactic and semantic status of a linguistic unit is altered to become less lexical and more grammatical.

This study presents evidence from medieval texts to claim that in the evolution from CL to OSp the demonstrative pronoun turned into a third-person pronoun. Furthermore, the study presents evidence from well attested MSp colloquial phenomena such as 'laísmo', 'leísmo' and indirect object doubling to suggest that in the evolution from MSp to FSp the clitic pronouns will become agreement markers.

Ultimately, this account proposes that the emergence of a novel category is not the product of the deterioration of an erstwhile one, and that novel categories are born from the interrelationship between the cognitive domain and the linguistic domain of the human mind.

Endnotes

1. It is not clear from the document in MP whether illos belongs to the preceding or the following predicate; thus the two possible predicates are shown.

2. This study assumes that the paradigms in (1) and (2) can all be defined and differentiated in terms of the features (or properties) defined below. The features are based on the following morphosyntactic description of CL ille: phonologically and syntactically free ([+ton]) unit which conveys definite ([+def]) and distal ([+dis]) reference ([+dei]) and which is morphologically marked for case ([+cas]).

- [deictic]: a paradigm is [+dei] if its forms point to (as opposed to qualifying or quantifying) a referent which is either spatial, textual, or grammatical. I will refer to these three functions as spatial, textual or discourse, and grammatical reference (or deixis), respectively. A category is [-dei] if it does not function in any of those three ways. Aquella in aquella casa 'that house' is [+dei], but mucha in mucha agua 'a lot of water' is [-dei].

- [definite]: a paradigm is [+def] if its forms refer to one specific entity or set of entities usually found in the surrounding context, i.e. a [+def] unit has a very narrow scope. Mis in mis hijos 'my sons' is [+def], but algunos in algunos hijos 'some sons' is [-def].

- [case]: a paradigm is [+cas] if its forms are morphologically marked for case, and [-cas] if they are not.

- [tonic]: a paradigm is [+ton] if its forms do not depend on any other word for prosodic stress and semantic interpretation, [+/-ton] if it shows tonicity allomorphy, and [-ton] if all the forms are dependent on other words.

- [distal]: a paradigm is [+dis] if it marks relative distance from the interlocutors, [+/-dis] if it has a focal or non-focal sense which does not mark relative distance, and [-dis] if none of its senses, focal and non-focal, mark relative distance.

3. An analogous process of reinforcement with the adjectival VL forms of ille gave rise to the new adjectival demonstratives such as aquela in the sentence below.

en aquela heredat
en aquella heredad
'in that estate'
G&H 13:22:14

In MSp the adjectival and pronominal forms of the new demonstratives are differentiated by means of a written accent over the latter.

4. The el cual antecedent pronouns derive analogously to the el que forms. The el que forms are more grammaticalized than the el cual forms because they are more commonly used in the spoken language, and thus have been generalized to more contexts as follows:

1) el cual and el que are interchangeable in oblique relative syntagms, but el cual has learned nuances as shown by the following sentence.

El hombre con el que hablé es de Grecia
 El hombre con el cual hablé es de Grecia
 'The man with whom I talked is from Greece'

2) el cual cannot be used in non-oblique relative syntagms which are the most common relatives as shown by the following sentence.

Compré los que él quería
 *Compré los cuales él quería
 'I bought the ones he wanted'

5. Although the declension of ille in (1) is subdivided into six subparadigms according to the threefold gender and the twofold number oppositions, (26) only shows four source paradigms given that both neuter subparadigms in (1) merged with the other four in the phase from CL to VL as follows. The neuter singular subparadigm blended with the singular masculine and the neuter plural subparadigm blended with the feminine singular as happened with other nominals. For example, the masculine singular MSp nouns tiempo 'time' and cuerpo 'body' derived from the neuter singular CL nouns tempus and corpus. Likewise, the feminine singular MSp nouns hoja 'leaf' and deuda 'debt' derived from the neuter plural CL nouns folia and debita.

6. A similar phenomenon of object doubling, with both direct and indirect object pronouns, is attested in the Pied Noir MFr dialect according to Roberge (1990).

References

- Díaz y Díaz, Manuel C. 1950. *Antología del Latín Vulgar*. Madrid: Gredos.
- Ford, Jeremiah Denis Matthias. 1934. *Old Spanish readings: selected on the basis of critically edited texts*. Boston: Ginn Company.
- Gifford, Douglas, and F.W. Hodcroft. 1966. *Textos lingüísticos del medioevo español*. Oxford: The Dolphin Book Co.
- Heine, Bernd, and Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.
- Heine, Bernd, and Ulrike Claudi, and Friederike Hünemeyer. 1991. *Grammaticalization: a conceptual framework*. Chicago: Univ. of Chicago Press.
- Lope Blanch, Juan M. 1990. *El español hablado en el suroeste de los Estados Unidos: Materiales para su estudio*. Univ. Nacional Autónoma de México: México, D.F.
- Menéndez Pidal, Ramón. 1956. *Orígenes del español*. Madrid: Espasa-Calpe.
- Roberge, Yves. 1990. *The syntactic recoverability of null arguments*. Québec: McGill-Queen's Univ. Press.

The 'double-go-future' construction in Spanish: A new grammaticalization phase?

Inés Miranda
University of Washington

I.- Grammaticalization: An overview

In their book *The evolution of grammar*, Bybee et al. 1994 (henceforth referred to as Bybee) set out to develop a theory of grammaticalization, a process by which "grammatical morphemes develop gradually out of lexical morphemes or combinations of lexical morphemes with lexical or grammatical morphemes" (Bybee:4). Their findings are a result of a cross-linguistic analysis from a diachronic perspective of the grammaticalization processes of morphemes associated with verbs, focusing on the way tense, aspect, and modality are molded into grammatical meaning.

Grammaticalization is seen as a single continuum of semantic generalization paralleled by a process of phonological reduction as grammatical morphemes or *grams* generalize their meaning and are used in a wider range of contexts, making them more dependent on surrounding material, both semantically and phonologically. This all leads to a rigidification of the syntactic position of a gram, which in turn favors a fusion process with other morphemes to which they have become associated. From this it follows that grams tend to be unbound or periphrastic early in the continuum, and bound or affixal in the later stages. Under this view, the dichotomy between analytic and synthetic ceases to be important, since the focus is on the beginning and end of the grammaticalization process, rather than on the point where a structure turns from analytic to synthetic. Along the same lines, the linguistic division between diachrony and synchrony is diminished given that even when a specific period of time is abstracted from the continuum (synchronic), it is still seen as a stage in the diachronic evolution path.

They claim that this process is *unidirectional*, that is, the change is always from semantic and phonological independence to dependence; once a gram becomes bounded, it will not become free again.

Central to their theory of grammaticalization is the claim that the developmental

In this paper, I will apply their analysis of the evolution of future grams to explain the synchronic stage of the Spanish 'go-future' *ir a + inf.* construction. I will claim that a new phase of grammaticalization in which a periphrastic construction consisting of a 'double-go-future' *ir a + ir a + inf.* (similar to the English 'gonna go' construction) is under way. This 'double-go-future' is being used colloquially to signal immediate future, a departure from its original meaning as the future of the verb *ir* 'to go'¹.

I will further claim that this 'double-go-future' construction is undergoing phonological reduction as a result of the convergence of the phonological constraints in the language regarding syllable structure and the semantic erosion of the preposition *A* 'to' as a marker of directionality towards a goal.

II.- Developmental path proposed for future grams in Bybee

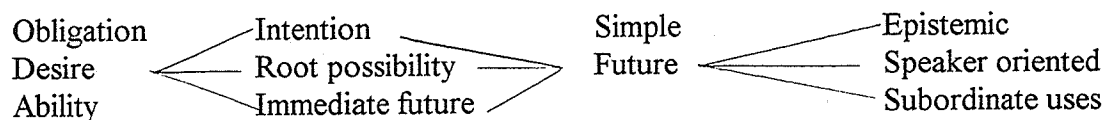
The future use of a gram is defined as "equivalent to a prediction on the part of the speaker that the situation in the proposition, which refers to an event taking place after the moment of speech, will hold" (Bybee:244). In their cross-linguistic data, forty nine out of the seventy languages with future grams had two or more forms. It is not surprising then to find a similar situation in Spanish where an older, simple or synthetic future form *cantaré* co-exists with a more recent periphrastic 'go-future' construction *voy a cantar*. This duplication of forms is explained by the *layer hypothesis* according to which the "independent development of grams from distinct sources and from similar sources at different periods, which produces layers of relatively old markers underlying layers of

¹ My analysis is based primarily on my own judgements as a native speaker and on those of other native speakers who were unfortunate enough to be my friends and had to stand long hours of close scrutiny and questioning. It should be mentioned that the others consulted and I come from Latin America, thus the present analysis might not reflect the situation of the Peninsular Spanish dialects. Special thanks go to Jeff Barnhart, Julia Mentan, Margaret Salome, Karyn Schell, and Bridget Yaden for their valuable insights regarding the 'go-future' construction in English, and to Lucía Menzinger for her useful comments during the presentation of this

more recently evolved ones" (*op.cit.*: 243). Given the layer hypothesis and the unidirectionality principle of the grammaticalization process, it follows that two or more grams might have similar uses along the future path, but that older grams reach, so to say, late-developing uses first. This is exactly the situation in Spanish as will be discussed below.

The lexical sources given for future grams are agent-oriented modalities², temporal adverbs and movement verbs, the latter being the most common source cross-linguistically.

The semantic development they propose for future grams can be seen in the following schema:



In the sections to follow we will see how the development of the Spanish future form *cantaré* from the Latin periphrastic future *cantare habeo*, as well as the rise of the 'go-future' and the 'double-go-future' constructions fit into this schema.

III.- *Cantaré* from *Cantare habeo*

Although the evidence from written texts is not conclusive as to the beginnings of the *habeo* construction as a future, it is thought that it was probably generalized in the eighth century. It co-existed with the older Classical Latin synthetic form *cantabo* and it still expressed obligation and sometimes possessive meaning from its lexical source *habere* 'to have' (see Fleischman 1982, hereafter referred to as Fleischman).

Fleischman argues that there was a semantic shift from expressing possession with *habere* to a nuance of obligation first, and later to a strong obligation to be carried out at

²

Agent-oriented modalities here include obligation, desire, ability, root

some posterior or future moment. Eventually, the meaning of possession and the nuance of obligation were lost:

possession -----> obligation -----> future
obligation

Bybee et al. on the other hand, consider that 'be' and 'have' constructions might have developed into futures from two sources: from possession > obligation, and from possession > predestination, given the arguments of Beneviste 1968 (cited by Bybee:261) with respect to the *habere* + *inf.* construction as not expressing obligation at all, but rather a sense of predestination that something 'will be done'.

In any case, the evolution from *cantare habeo* reflects the evolution either from obligation or predestination to intention first, and ultimately to future. The important fact to consider here is that both origins, be it obligation or predestination, converge later into expressing intention, an intention projected into the future.

The new synthetic future *cantaré* represents a later stage in the grammaticalization path of the old periphrastic construction *cantare habeo*. The complete development can be seen in the following table adapted for Spanish from Fleischman:

Morphosyntactic evolution of *cantare habeo* (adapted from Fleischman 1982:71)

<u>Stage</u>	<u>Form</u>	<u>Level</u>	<u>Meaning</u>
Classical Latin	cantare habeo [inf.] [main vb]	Syntax	possession 'I have [something] to sing'
Late Latin	cant a re habeo [main vb] [aux]	Syntax	obligation-future 'I have to sing' - 'I will sing'
Romance	cant a r é [stem] [tense mood] [person number]	Morphology	future 'I will sing'

The shift from the Classical Latin SOV word order to SVO was a determinant factor in the fusion of the periphrastic future form *cantare + habere*. The order verb + aux. reflected the older SOV pattern, making the structure more vulnerable to change. This order provided for the retention of the morphological person-number markers in the suffixed inflection in the synthetic future form, as can be seen in the table above. As Fleischman points out, the opposite order *habere + cantare* would either caused the loss of this morphological information altogether or produced anomalous prefixed or infixes inflections:

hab-eo cant- are	(a) */akantár/
	(B) */okantár/, */askantár/, */akantár/

(Fleischman:115)

Retention of important morphological information thus seems to play a crucial role in the degree of fusion that a gram might undergo, a fact that follows from the semantic-phonetic parallel reduction hypothesis proposed by Bybee: phonetic reduction goes on a par with semantic reduction (see also discussion in VI below).

In Spanish, the new synthetic future co-existed with the older *habeo* construction until the seventeenth century and eventually supplanted it in all its uses.

IV.- The 'go-future' construction *Ir a + inf.*

As mentioned earlier, movement verbs constitute the most common lexical source for future grams. The meaning of 'movement toward' of lexical verbs such as *ir a* 'to go to' which implies both a spatial and a temporal movement, expresses also an intention:

- | | |
|--|---|
| 1a. Vamos al cine.
(We) go to the movies
'Let's go to the movies.' | 1b. Voy al museo.
(I) go to the museum.
'I am going to the museum.' |
|--|---|

As the meaning of intention expressed by the verb *ir* became more generalized, the idea of a temporal movement rather than a spatial one was reinforced. Therefore, the

semantic shift involved in this grammaticalization process is one of loss of spatial meaning: the subject fulfills an intention without moving spatially:

2. Voy a estudiar mañana.
 (I) go to study tomorrow
 'I am going to study tomorrow.' - 'I will study tomorrow.'

It is interesting to note that if the construction is intended to mean futurity, that is, involving only a temporal movement, no temporal adverbs are allowed to interfere in between:

3. Vamos a cantar mañana.
 (We)go to sing tomorrow
 'We are going to sing tomorrow.' - 'We will sing tomorrow.'

If an adverb is placed in between, the verb *ir* retains its lexical meaning as expressing spatial movement:

4. Vamos mañana a cantar.
 (We)go tomorrow to sing
 'We are going [somewhere] to sing tomorrow.'
 *'We will sing tomorrow.'

This is consistent with the hypothesis that as grams grammaticalize, they become more dependent on surrounding material, both semantically and phonologically, and their syntactic position is rigidified.

Bybee points out that in order to derive future, an allative component meaning 'toward' must be present either in the verb itself or in the construction. This directionality sense is expressed by the preposition *A* in all Western Romance languages with actual movement verbs (implying spatial movement). Thus, a sentence such as 'I go to my friend's house' is translated as the verb 'to go' followed by the preposition *A*:

'I go to my friend's house'

- Port. Vou a casa do meu amigo.
 Spa. Voy a la casa de mi amigo.
 Ital. Vado a casa di mio amico.
 French Je vais à la maison de mon ami.

Here, the preposition indicates a terminal point for the action of moving, which contrasts with the use of the preposition *de* 'from' to indicate a point of departure. This use might be better expressed in Spanish by the preposition *hasta* 'until', as Gili Gaya states:

Cuando se trata de distancias o de medidas de tiempo, el sentido local o temporal se combina con la idea de movimiento [...] El punto de partida está bien determinado por la preposición *de*; el punto terminal del espacio o del tiempo va marcado por la preposición *a*, pero lo estaría con más precisión por *hasta*.
(Gili Gaya 1990:250)

Therefore, let us conclude that all these languages share the use of the preposition *A* as an equivalent to *hasta* when a spatial movement is inferred, that is, the use of *A* to signal the end point or destiny of movement.

However, only Spanish uses *A* with the more grammaticalized 'go-future' construction, in which the movement verb has lost its spatial reference, as can be seen in the translation of the sentence 'I am going to sleep'³:

"I am going to sleep" - "I will sleep"

Spa.	Voy a dormir.
Port.	Vou dormir.
French	Je vais dormir.

This was not always the case in Spanish for constructions with and without the preposition are both attested in Medieval texts (see Cuervo 1886). One cannot but speculate that the preposition *A* in Spanish eventually became reanalyzed to signal a goal without necessarily implying actual spatial movement. In other words, the preposition *A* underwent a grammaticalization of its own in Spanish, unlike in the other Romance languages.

We have seen how the 'go-future' construction evolved from the lexical verb *ir*, implying spatial as well as temporal movement, to the periphrastic construction *ir a + inf.* with future use through a semantic generalization of its original meaning consisting of the loss of spatial reference.

In Modern Spanish, the 'go-future' construction co-exists with the older inflectional future form *cantaré*. In Latin American dialects this older form is restricted to the late-developing epistemic use, in correspondence with the unidirectionality hypothesis and the developmental path for future grams proposed in Bybee:

- 7a. Parecerá tonto, pero no lo es.
 (He) seems (fut.) dumb, but not it is
 'He might look dumb, but he is not.'
- 7b. Va a parecer tonto, pero no lo es.
 (He) goes to seem dumb, but not it is
 'He will look dumb, but he is not' (future reading only).

Although the 'go-future' is the predominant, almost exclusive form to express simple future⁴, a subtle contrast might still obtain between the inflectional and the 'go-future' construction when used to express simple future:

- 8a. Me vas a extrañar.
 me (you) go to miss me
 'You are going to miss me.'
- 8b. Me extrañarás.
 me (you) miss (fut.)
 'You will miss me.'

8b, with the older inflectional future has a stronger sense of obligation-predestination than 8a, a contrast that also seems to hold for English. However, I would not yet consider 8b to be epistemic since it expresses neither possibility nor the degree of commitment of the subject to accept the truth value of the proposition⁵. This contrast reflects that the older future gram has remnants from its source construction *cantare habeo* that expressed such early uses as obligation-predestination.

V.- The 'double-go-future' construction *Ir a + ir a + inf.* A new phase of grammaticalization?

The 'go-future' construction is naturally used in combination with the infinitive *ir* 'to go' to express its future form as a lexical verb of moving:

⁴ The older future form is only retained as a simple future in extremely formal written language, particularly in order to avoid repetition.

⁵ I am indebted to Professor Jurgen Klausenburger for pointing out that 8b might only be expressing the now exclusive epistemic use of the older

9. Voy a ir a la piscina a nadar.
 (I) go to go to the pool to swim
 'I'm going to go (gonna go) to the pool to swim.' - 'I will go to the pool to swim.'

This construction obviously expresses spatial as well as temporal movement, but when this construction is used in contexts where the actual spatial distance is irrelevant, it acquires an immediate future meaning:

- | | |
|--|---|
| <p>10a. Voy a ir a bañarme.
 (I) go to go to take a shower
 'I'm (going to go)gonna go take a shower.'</p> | <p>10b. Voy a bañarme
 (I) go to take a shower
 'I'm (going to) gonna take a shower.'</p> |
|--|---|

10a. expresses the intention of the subject to take a shower right now, as opposed to some time in the future, as expressed by 10b⁶. Let us contrast these examples with 11a and b in which a future temporal adverb has been added:

- 11a. Voy a ir a bañarme el viernes.
 (I) go to go to take a shower on Friday
 'I'm going to go [somewhere] to take a shower on Friday.'
- 11b. Voy a bañarme el viernes.
 (I) go to take a shower on Friday
 'I'm going to (gonna) take a shower on Friday.'

In 11a the temporal adverb places the action outside the immediate future, and thus the 'double-go-future' construction is understood as the specific future use of the lexical verb 'to go'. It cannot have the same meaning as in 10a. On the other hand, sentence 11b shows the use of the 'go-future' construction as a simple future. Consequently, we can conclude that the periphrastic future of the verb 'to go' has undergone a similar process of semantic reduction by which the spatial-temporal reference expressed by the infinitive 'to go' has given way to express intention in the immediate future, one of the earlier uses of future grams.

Analogous to sentences 3 and 4 in section IV, we can see a more fixed position and a greater dependence among the elements of the new grammaticalized 'double-go-future' construction than among those of the 'go-future' of the lexical verb 'to go':

12a. Voy a ir a bañarme ahora.
 (I) go to go to take a shower now
 'I'm going to go (gonna go) take a shower now.'

12b. Voy a ir ahora a bañarme.
 (I) go to go now to take a shower
 'I'm going [somewhere] now to take a shower.'

The only possible interpretation when there is intervening material is the one expressed in 12b, that is, the grammaticalized 'double-go-future' construction in 12a which means immediate future is more bound as predicted in Bybee and does not allow intervening material.

Thus, in Modern Spanish, we are witnessing the rise of a new periphrastic construction stemming from a verb meaning 'to go' to express immediate future, something that appears to be totally innovative⁷.

VI.- Further phonological reduction of the 'double-go-future' construction

In section IV above (see pp. 6-8), the status of the preposition *A* as a marker of directionality, end point or goal in Western Romance languages was discussed, particularly the contrast between its absence or presence in the 'go-future' constructions. Presumably, Spanish, the only Western Romance language that retains it in the 'go-future' construction does so because it has generalized its use as a marker of direction, goal or objective even when there is no actual spatial movement involved. Thus, the preposition

⁷ Bybee reports that in their cross-linguistic data, none of the languages with

in the 'go-future' construction indicates the objective or goal predicted to occur in the temporal future, rather than a terminal end point of a movement verb:

13. Voy a trabajar mañana.
 (I) go to work tomorrow
 'I will work tomorrow.'

In 13, the action of working is the goal predicted to take place in the future. Accordingly, it is marked by the preposition *A*.

Similarly, in the 'double-go-future' construction, the goal predicted to occur in the immediate future is the one expressed by the infinitive following the infinitive *ir* 'to go', and not the verb *ir* itself:

14. Voy a ir a vestirme.
 (I) go to go to get dressed
 'I'm going to go (gonna go) get dressed.'

The goal expected to be accomplished in the immediate future is to get dressed, not the action of 'going' expressed by *ir*, whose original meaning as a verb of movement has been grammaticalized here to convey the notion of immediateness in the future. This goal is marked by the preposition *A* following the infinitive *ir*.

The same can be said of the plain 'go-future' construction used as the simple future of the lexical verb *ir* 'to go'; the ultimate goal in this case is the terminal spatial locative in the construction:

15. Voy a ir a la biblioteca.
 (I) go to go to the library
 'I will go to the library.'

The preposition *A* which precedes the infinitive *ir* in 14 and 15 has no semantic content. Moreover, given the tendency in languages like Spanish to avoid onsetless syllables (see Kenstowicz 1994, especially chapter 6), this preposition stands out in the language as the only one not conforming to this onset principle. The convergence of the semantic erosion and the phonology of the language make the preposition *A* prone to deletion in constructions like sentences 14 and 15, a process that has already occurred in certain Latin American dialects, particularly in Central America. Had it not been for the

fact that the preposition *A* has undergone a process of semantic bleaching in the grammaticalized construction, the phonological reduction process would have been unlikely to happen if not implausible, a fact that supports the semantic-phonological parallel reduction argued for in Bybee. Note for example that *A* is not deleted in the contracted form of the 'go-future' construction for first person singular which is heard in some Central and South American dialects, given that it **does** signal the action or objective to be accomplished in the future, as discussed above:

16. Voy a darte un regalo bonito. = [bwa] darte un regalo bonito.
 (I) go to give you a nice present
 'I will give you a nice present.' - 'I'm going to (gonna) give you a nice present.'⁸

Indeed, this preposition would not exist at all in the language if phonology were to supersede morphology at all times. Its morphological functions have so far guaranteed its existence in the language. This is not to say that morphological markers are invariably 'protected' against phonologically motivated processes (witness the loss of final consonants in French), but that the tendency is more towards the retention of morphological information, a tendency that has played an important, sometimes explanatory role in sound change and language change in general (i.e., retention of final /s/ only in those Romance languages where it is a marker of plurality; see Harris 1988:37. Also, see discussion of the grammaticalization of *cantaré* from *cantare habeo* in section III.) The parallel reduction hypothesis constitutes, therefore, a well-founded formalization of these observations.

A different approach to the grammaticalization of the Spanish future forms is proposed in Anderson 1979. Anderson argued for a new phase of grammaticalization of the 'go-future' construction in Spanish through a process which he called Morphologization II, by which "phonological rules give rise to a fusion of lexemes which create a new morphology" (Anderson 1979:31).

According to him, the preposition *A* in the 'go-future' construction has undergone syncope, and the form of the verb *ir* has generalized to the third person singular *va* in the

⁸ The first person singular is the only contracted form of *ir* + *a* in the 'go-

Panamanian dialect, so that *va* is predicted to become a prefix which will be the morphological marker of future tense: *vadormir* 'I will sleep, you will sleep, he/she will sleep'. To my knowledge, no such forms are attested in any dialect, and if they are, I strongly disagree with Anderson's prediction for two reasons:

- a) The person-number markers would be lost completely, a situation that would be extremely exceptional in the Spanish verbal system. This would force Spanish, a pro-drop language, to recur to the obligatory use of subject pronouns.
- b) Such a prefixation process is disfavored in general given psycho-linguistic perception and processing factors: suffixation and prefixation are not symmetrical in terms of perception and processing since in prefixation, both, the stem and the prefix have to be analyzed by the listener, whereas, in suffixation, only the stem needs to be analyzed (Hall 1992, cited by Klausenburger 1996).

Along these lines, Klausenburger 1996 maintains that the typology of languages (right vs. left branching) might be an important factor in determining the degree of grammaticalization that is likely to happen: right-branching languages, such as Spanish, have a 'built-in' mechanism against advanced grammaticalization given the position of the elements in the sentence and the prefixation dispreference already mentioned.

On the other hand, the deletion of the semantically null preposition *A* preceding the infinitive *ir* in the 'double-go-future' construction goes along the lines of the parallel reduction hypothesis and leads to a suffixation process internal to the grammaticalized construction: Once the preposition disappears, the infinitive *ir* is fused with the conjugated form of the verb 'to go' to conform to the onset principle:

Fused forms

voy a ir	>	voy ir	>	vuir [buyr]	'I am going (gonna) to go'
vas a ir	>	vas ir	[ba.sir]		'You (sing.) are going to go'
va a ir	>	va ir	[bayr]		'He/she is going to go'
vamos a ir	>	vamos ir	[ba.mo.sir]		'We are going to go'
van a ir	>	van ir	[ba.nir]		'They are going to go'

Morphologically speaking, there is no information lost because the person-number markers are retained in the new fused forms, and the *ir* suffix is analyzed as the future-tense marker.

Could these be the new forms of a synthetic-inflectional future? Synchronically, these forms are not very frequent in the language given that they are restricted to the 'go-future' of the verb 'to go' and to the 'double-go-future' construction which expresses immediate future, still in an incipient stage. However, if the latter is to continue its path along the grammaticalization line, then these forms may as well develop as a new inflectional future in Spanish.

VII.- Conclusions

The diachronic evolution and the synchronic status of future grams in Spanish corresponds to the grammaticalization path for future grams proposed in Bybee et al. 1994.

Synchronically, two future grams co-exist in the language: an older synthetic form *cantaré* which evolved from the Latin periphrastic future *cantare habeo* and eventually supplanted it, and a more recent periphrastic 'go-future' construction *ir a + infinitive*, *voy a cantar*. This co-existence of grams is explained by the layering hypothesis, and their uses, by the unidirectionality principle: The older gram has been eventually supplanted by the new periphrastic form in the early uses, being restricted now to the late-developing epistemic use, while the 'go-future' construction expresses simple future.

In some Latin American dialects, the 'go-future' construction of the lexical verb *ir* 'to go' is undergoing grammaticalization, becoming a new form to express the more specific meaning of immediate future. I have called this form the 'double-go-future' construction, *ir a + ir a + infinitive*. It is claimed here that this new construction is undergoing a process of phonological reduction parallel to a process of semantic bleaching as the result of its grammaticalization. This twofold reduction has led to the fusion of the conjugated form of the construction with the infinitive *ir*, which, in the future, might become new forms of an inflected future tense in Spanish.

The synchronic stage of future grams in these Latin American dialects can be represented as follows:

Immediate future

Simple future

Epistemic



synthetic cantaré

ir a + infinitive

ir a + ir a + infinitive

References

- Anderson, Eric W. 1979. The development of the Romance future tense: Morphologization II and a tendency toward analyticity. *Papers in Romance I*: 21-35.
- Bybee, Joan et al. 1994. *The evolution of grammar. Tense, aspect, and modality in the languages of the world.* Chicago and London: University of Chicago.
- Cuervo, Justo Rufino. 1886. *Diccionario de construcción y régimen de la lengua castellana. Tomo primero.* París: Libreros editores.
- Fleischman, Suzanne. 1982. *The future in thought and language.* Cambridge: CUP.
- Gili Gaya, Samuel. 1990. *Curso superior de sintaxis española. 15th. ed.* Barcelona: Bibliograf, S.A.
- Harris, Martin and Vincent Nigel (eds.). 1988. *The Romance languages.* New York: Oxford University Press.
- Kenstowicz, Michael. 1994. *Phonology in generative grammar.* Cambridge, Mass.: Blackwell.
- Klausenburgen, Jurgen. 1996. *Seminar on Grammaticalization.* University of Washington. Lecture notes.

- Rizzi, L. (1990). *Relativized Minimality*. Cambridge, Massachusetts: MIT Press.
- Saito, M. (1985). *Some Asymmetries in Japanese and Their Theoretical Implication*. Ph.D. Dissertation, MIT.
- _____ (1992). Long Distance Scrambling in Japanese. *Journal of East Asian Linguistics* 1: 69-118.
- Suñer, M. (1991). Indirect Questions and the Structure of CP: Some Consequences. In H. Campos and F. Martinez-Gil eds., *Current Studies in Spanish Linguistics*. Washington, DC: Georgetown University Press.
- _____ (1993). About Indirect Questions and Semi-Questions. *Linguistics and Philosophy* 16: 45-77.
- Watanabe, A. (1991). Wh-in-situ, Subjacency, and Chain Formation. ms., MIT.
- _____ (1993). *AGR-Based Case Theory and Its Interaction with the A-Bar System*. Ph.D. Dissertation, MIT.
- Whitman, J. (1989). Topic, Modality, and IP Structure. *Harvard Studies in Korean Linguistics III*. Seoul: Hanshin Publishing Co.
- Wilder, C., and D. Cavar (1994). Word Order Variation, Verb Movement, and Economy Principles. *Studia Linguistica* 48-1: 46-86.
- Yang, D.-W. (1989). *Jibay-Kyelsok Ilonyi Kicho*. [An Introduction to Government and Binding Theory]. Seoul: Shin A Sa.
- _____ (1993). The Minimalist Theory and the Structure of Korean. presented to the 2nd Seoul International Conference on Generative Grammar.
- Yoon, J.-Y. (1990). *Korean Syntax and Generalized X-Bar Theory*. Ph.D. Dissertation, University of Texas at Austin.

LIST OF CONTRIBUTORS

José Carrasquel
1798 Scenic Ave. Box 321
Berkeley CA94709

Kyu Hwang
kyu@u.washington.edu

Yongkil Jeon
yonkil@u.washington.edu

Tai-Soo Kim
taisoo@u.washington.edu
AF ACAD Cheongwon
P. O. Box 2, Chungbuk, KOREA 363-849

Inés Miranda
ines@u.washington.edu

Alice Taff
taff@u.washington.edu

WORKING PAPERS IN
LINGUISTICS
UNIVERSITY OF WASHINGTON
VOLUME 13 **JULY 1996**

TABLE OF CONTENTS

Acoustic Correlates of Dec Xinag Vowels Alice Taff	1
Licensing and Scope Interpretation of Wh-Phrases in Korean Tai-Soo Kim	33
A Case-Tier Approach to Finnish: Feature-based Definition of Case Domain Yongkil Jeong	75
The English Subjunctive and Minimalist Case Checking Theory Kyu-Hong Hwang	95
The Evolution of Pronominal <i>ille</i> from Latin to Spanish: A Grammaticalization Analysis José Carrasquel	121
The 'double-go-future' construction in Spanish: A New Grammaticalization phase? Inés Miranda	147