UNIVERSALS OF WORD FORMATION PROCESSES: NOUN INCORPORATION IN THE ACQUISITION OF SAMOAN AS A SECOND LANGUAGE[1]

PETER J. ROBINSON
University of Hawai'i

Studies of second language syntactic acquisition (e.g., Gass 1979; Flynn 1991; White 1989) have been motivated by research into language universals within both typological (Greenberg 1966), and universal grammar (Chomsky 1981) frameworks. In the typological framework learning difficulty is predicted by the relative markedness of the second language, when compared with the learner's first language (Eckman 1985). In the universal grammar framework learning difficulty is predicted where there are differences in the parameter settings of the first and second language (Flynn 1989). Some support exists for the influence of markedness on learning difficulty (e.g., Gass 1979; Pavesi 1986; Doughty 1991), but there is disagreement about the influence exerted by universal grammar (Bley-Vroman & Chaudron 1990; Flynn & Lust 1990). To date no studies of second language word formation processes have been motivated by language universals research from either of these perspectives. This study examines the influence of a proposed implicational hierarchy (Mithun 1984) and constraints of universal grammar (Baker 1988) on the acquisition of noun incorporation processes by second language learners of Sāmoan. The methodology involved reaction time, grammaticality judgement and response certainty measures of the processing difficulty and acceptability of examples of noun incorporation for English speaking learners of Sāmoan, with the latter measure giving the clearest support for two hypothesized orders of difficulty.

INTRODUCTION

1.1 First and second language research on the acquisition of word formation

Developmental studies of the acquisition of word formation processes in first language acquisition (e.g., Clark 1981,1982; Clark & Berman 1984; Downing 1977) have been based on descriptive classifications (e.g., Adams 1973;

[1]I am grateful to William O'Grady of the Dept. of Linguistics at the University of Hawai'i, who encouraged this project as a term paper for his course on syntactic analysis, and to the following for their comments; Dean Mellow, University of British Columbia; Robert Bley-Vroman, Dept. of ESL, and John Mayer, Dept. of Indo Pacific Languages- both at the University of Hawai'i . Versions of this paper were presented at the AAAL conference, March 1992,and SLRF, April 1992.

Marchand 1969) that have had little to say about the question of markedness and universals from either a typological (e.g., Greenberg 1966; Hawkins 1985) or a Chomskyan perspective (Chomsky 1981). However, at the level of syntax and phrase structure, notions of markedness and universals have been central to the issue of learnability (e.g., Pinker 1989) and speculation about the mechanisms necessary for acquiring knowledge of language(e.g., Berwick 1985; Williams 1981).

No doubt this is partly a reflection of the fact that , to date, little work has been done in descriptive or theoretical linguistics that could motivate universal–theoretic approaches to word formation acquisition (though this position is now changing, see Baker 1988; DiScullio & Williams 1989). Instead, explanations for the emergence of word formation abilities in child language have been motivated either by appeal to the general processing complexity of the operations involved in interpreting and producing new forms, or to input factors such as frequency of use of a word formation device in parental speech.

Specifically, explanation of the degree of difficulty of acquiring L1 word formation patterns has focussed on i.) the relative simplicity of different word–forms, ii) their semantic transparency of meaning and iii) their productivity or frequency of use. (Clark 1981, 1983). For example, 'semantic transparency' (the principle that known elements with one to one form meaning mappings will be preferred in forming new words) accurately predicts , in child first language development, that new nouns will initially be formed by compounding already known nouns (Downing 1977). At a later stage word formation processes make use of potentially ambiguous morphemes like agentive 'er' only after it has been functionally distinguished from comparative 'er'. Children also seem to progressively complexify the innovations they produce, making as few 'formal' changes as possible in new word forms–to the extent of keeping the base forms of new words intact (e.g., saying 'pompous–ity', instead of 'pompos–ity' , or 'electrik–ity' instead of 'electric–ity' see Romaine 1983, and Cutler 1980), and to favour those formation processes that appear to be most 'productive' in adult speech, for example using 'ness' to form nouns rather than less common 'ity' (see Aitchison 1987; Clark & Berman 1984).

There have been few studies of the second language acquisition of word

formation processes, but what studies there are (e.g., Olshtain 1987) also make no reference to notions of markedness and universals. The motivation for developmental predictions is basically quantitative, and the hypothesis is that increased passive recognition and productive use of novel word forms will increase with increased target language proficiency. In contrast, this paper examines evidence for the influence of i) a markedness hierarchy (Mithun 1984, 1986) and ii) universal grammar (Baker 1988) on the acquisition of one example of a word–formation process, that of noun incorporation (NI) by English speaking learners of Sāmoan as a second language.

1.2 Noun Incorporation and Implicational Universals

Noun incorporation is the process of moving a noun from an independent base structure position to combine with the verb in the surface syntax. True noun incorporation occurs in languages like Onondaga and Southern Tiwa where the resulting noun–verb compounds function as the main verbs of their clause. This is to be distinguished from the process of noun–verb compounding in English, e.g., 'Pat is a hopeless money–loser', where the resulting compounds are deverbal nouns (see Baker 1988: 78). Mithun (1984, 1986) has proposed a discourse–functional explanation for the evolution of noun incorporation across languages and identified a resulting implicational hierarchy.[2] Her claim is that once noun incorporation appears in a language it develops along an invariant path. Stage 1 involves lexical compounding in which a noun and verb combine to form an intransitive verb. Stage 2 allows an oblique argument to assume the syntactic role vacated by the incorporated noun. At stage 3 nouns reflecting known information are incorporated to narrow the scope of the verb, fulfilling a discourse organizing function, while finally in stage 4 a generic noun is incorporated to qualify the verb, along with a more specific NP that identifies the implied referent. Occurrence of stage 4 processes in a language implies occurrence of all the others.

[2]This is a true typological hierarchy, with language types corresponding to the stages discussed above. Mithun's claim is not a claim about acquisition stages, though it is the intention of this paper to explore the relationship between the typological hierarchy and L2 learning difficulty.

Each stage, then, in the evolution of noun incorporation in a language corresponds to the development of a particular type of incorporation. For example, in Stage 1 a verb and noun are combined to form an intransitive predicate, (type 1 incorporation). The incorporated noun no longer has referential status, and loses marking for number and definiteness.

1 a) Gu bea chuwqiy eamareaw
I PRES buy CONN copra
'I am buying copra'

b) Gu bea chuwaay'–mareaw
I PRES buy–copra
'I am copra–buying' (Yapese:from Mithun 1984: 850)

At stage 2 NI involves a change in case relations, as in the following example where in the unincorporated sentence a) the direct object is 'face', marked by '–yos–', whereas in b) the direct object, after incorporation, has changed to the owner of the face;

2 a) s–oßa a–yos–ey
his–face I–it–wash
'I washed his face'

b) a–s–oßa–ey
I–him–face–wash
'I face–washed him' (Tupinamba:from Mithun 1984: 857)

Stage 3 involves the same process of incorporation as in stages 1 and 2, but the motive and context for the incorporation is different. Here the object NP is incorporated for the purpose of backgrounding it as information . In the following example 'meat' is originally unincorporated, but subsequent reference to 'meat' is incorporated. This is taken to be an index of the speaker/writer's decision to reduce its prominence as known information in the discourse (see Givón 1985 and Haiman 1985 for similar analyses).

3. askeman ti–'–kwa nakatl. na' ipanima ni–naka–kwa.
never you–it–eat meat I always I–meat–eat.
(Nahuatl:from Mithun 1984: 860)

These three types of incorporation all involve altering the argument structure of the verb so that an internal argument is satisfied within the verb. Where the verb only takes one internal argument this results in a detransitivization of the verb (type 1). Where it takes two internal arguments then incorporation of direct object frees the oblique to receive accusative case (type 2). The motive for both of these types of incorporation can be discoursal so that a previously unincorporated NP is subsequently incorporated as a way of backgrounding already known information (type 3). However, in the fourth Stage of NI the incorporated noun does not satisfy an argument of the verb, it serves only to classify, semantically, the direct object NP. The classifier noun is a more general superordinate term, as in the following example:

4. bene–dulg–nan mangaralaljmayn
they.two–tree–saw cashew.nut
'They saw a cashew tree'

In this case the incorporated noun, 'tree', serves to classify the object NP 'cashew'. Mithun's hierarchy is thus type/stage 1 < 2 < 3 < 4.

1.3 Noun incorporation and universal grammar

1.3.1 Head movement

A different perspective on language universals is taken by linguists working within the Chomskyan framework of Universal Grammar. Recent

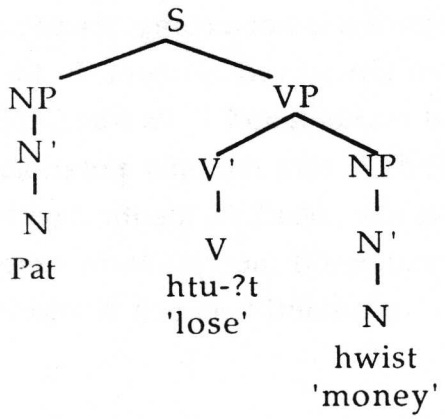
work in this framework has proposed that lexical processes like word formation are sensitive to the same restrictions governing the acceptability of larger syntactic structures of a language. In particular, Baker's claim (1988) is that incorporated forms are derived from deep structure via a syntactic process of head movement that conforms to the empty category principle, i.e. that a moved element must c–command its trace, and that there must be no barriers along the c–command path. For example, the following examples, from Onondaga (see Baker 1988: 76) are thematic paraphrases, with the verb assigning accusative case to 'money', but they differ in that in 5b) the direct object has been incorporated into the verb stem, whereas in 5a) it is a separate word.

5 a) Pet wa?–ha–htu–?t–a? ne? o–hwist–a?
Pat PAST–3MS/3N–lost–CAUS–ASP the PRE–money–SUF
'Pat lost the money'

b) Pet wa?–ha–hwist–ahtu–?t–a?
Pat PAST–3MS–money–lost–CAUS–ASP
'Pat lost money'

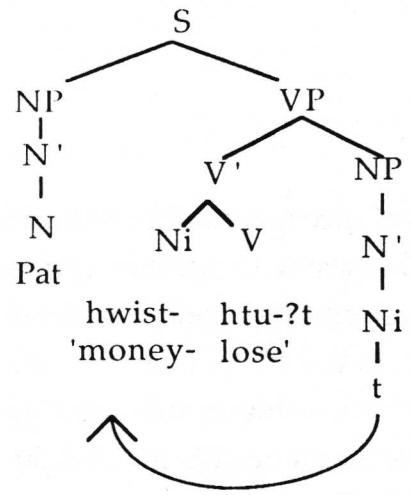
Since the two sentences display the same thematic assignment relationship then following the Uniformity of Theta Assignment Hypothesis, (that NPs receiving the same theta roles must occupy the same position in deep structure), the deep structure for both the sentences must be:

6)



For 5a), the unincorporated version, inflectional morphology is added to produce the surface form of the sentence, but for 5b) an additional movement adjoins the structurally lower NP to the verb. The Projection Principle states that movement cannot destroy thematically relevant structure, so the surface structure for is:

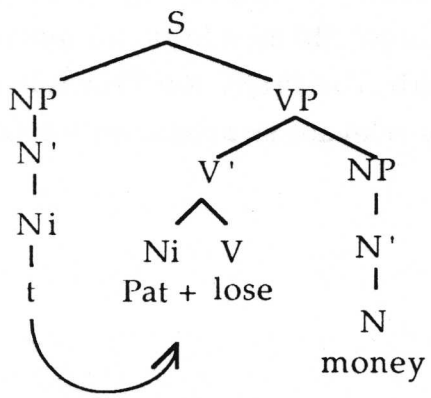
7)



For the surface structure to be well formed the movement must conform to the

ECP, i.e. that a trace must be properly governed. Proper government requires A and B to be co–indexed. A must also c–command B, and there must be no barriers on the c–command path. An XP is a barrier if it is not a complement of A, and if it has a head distinct from the head of the moved element. In the example above the ECP is satisfied. A, the moved element, and B, its trace, are coindexed by antecedent indexing. A c–commands B, and the only potential barrier to government, NP is a complement, and it has a head, N, identical with that of the moved element. The ECP would be violated if incorporation were from the subject NP as in this example, where c–command between A and B does not hold.

8) *



In the following section I introduce examples of incorporation in Sāmoan and locate them in relation a) to the implicational hierarchy of Mithun (1984), and b) to the principles of well formedness underlying proposals about Universal Grammar discussed above, i.e. conformity to the ECP. I then describe a study of the acquisition of Sāmoan word formation, using grammaticality judgements and reaction times as measures, and discuss the results in the light of the above, supposed, universal constraints on such processes.

NOUN INCORPORATION IN SAMOAN

2.1 Sāmoan word order

Sāmoan is a verb initial language though, 'the order of post verbal phrases is very free' (Hovdhaugen 1985: 86) with a preference for VSO in literary Sāmoan. Subject is usually marked with 'e', and object with zero.

'Sa	f	ai	e	Laufai	le	galuega'
	PST	did	SUBJ	Laufai	the	job

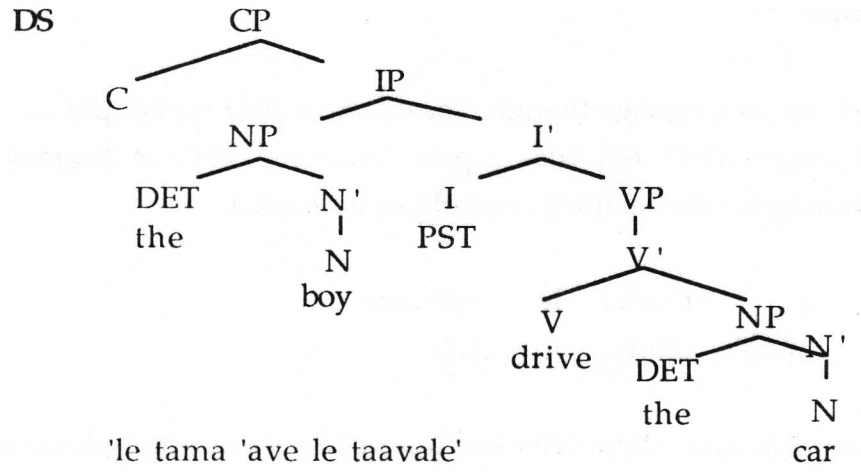
The verb can combine with up to three NPs and a number of temporal phrases whose position is not fixed. Sproat (1985) has proposed an analysis of verb initial languages in which they are derived from deep structure SVO order by movement of the verb to Infl, then to Comp. This analysis is adopted and illustrated in the following section.

2.2 .Examples of grammatical noun incorporation in Sāmoan

2.2.1 Lexical compounding

Mithun's first category of incorporation, detransitivizing noun compounding, is very productive in Sāmoan. It leaves an empty NP after incorporation giving the DS in 11) for example 13a) below.

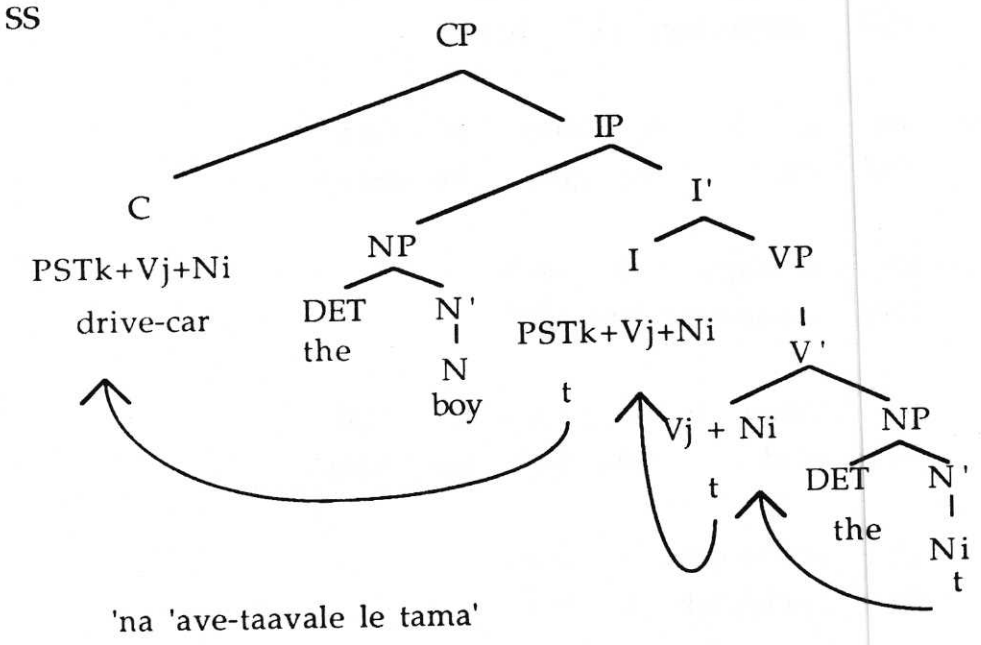
11)



The surface structure of the incorporated sentence in 13b) below is derived by a series of head movement operations, as shown in this structure.

First the noun 'car' is moved from its position as complement of V to attach to the verb (Chomsky-adjunction). Then V+N are moved up to Infl where they attach to 'PST' giving PST+v+n, and finally Infl is moved up to the empty C position. Each of these movements satisfies the ECP, so the traces are properly governed. The resulting order of morphemes in the final incorporated verb , PST–V–N reflects the sequence of movement operations via which the surface structure was derived, thus conforming to the 'mirror principle' (Baker 1988: 13). The same analysis applies to all the other examples of Sāmoan noun incorporation given in this section.

12)



'na 'ave-taavale le tama'

13 a) na 'ave e le tama le taavale
PST drive the boy the car

b) na 'ave–taavale le tama
PST drive–car the boy

14 a) na fana e le tama le lupe
PST shoot the boy the dove

b) na fana–lupe le tama
PST shoot–dove the boy

15 a) na inu e le tama le pia
PST drink the boy the beer

b) na inu–pia le tama
PST drink–beer the boy

16 a) na ai e le teine le mago
PST eat the girl the mango

b) na ai–mago le teine
PST eat–mango the girl

17 a) na faitau e le tama le tusi
PST read the boy the book

b) na faitau–tusi le tama
PST read–book the boy

18 a) sa fau e le tama le fale
PST build the boy the house

b) sa fau–falele tama
PST build–house the boy

2.2.2 Classifier incorporation

The occurrence of classifier incorporation (CI) (Mithun's type 4) in a language implies the occurrence of all the other forms of incorporation discussed previously. However CI is not characteristic of Sāmoan, all of the following examples having been judged unacceptable by native speakers.

*19. naia fualaau–fofoe le fai
PST he fruit–peeled the banana

*20. na la aina–ai le keke
PST he meal–ate a cake

*21. na inu–faafau le pia
PST she drink–bought the beer

*22. sa manu–faatau le maile
PST I animal–bought a dog

*23. sa laau–vaai i le mago
PST I tree–saw a mango

*24. sa ou ofu–su'i pulepule
PST I dress–made a polka dotted one

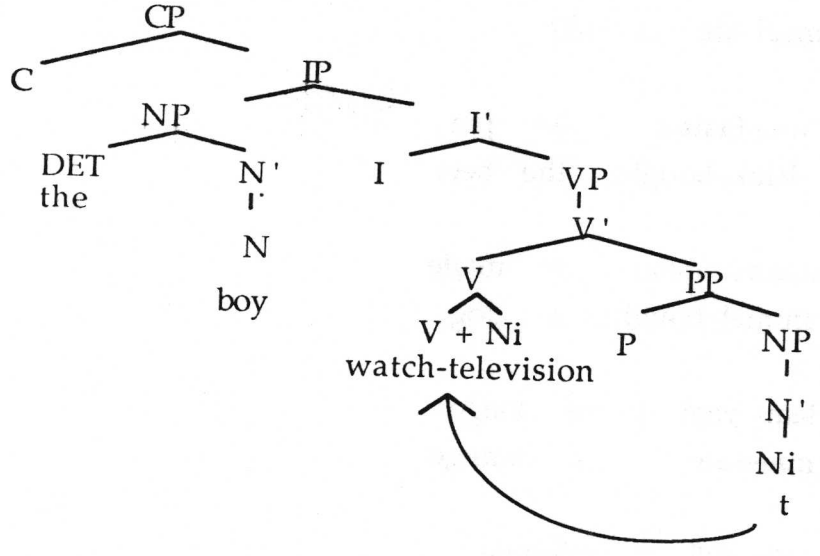
In section 3. I discuss the motivation for, and implications of, using examples of this type in grammaticality judgement and reaction time measures in light of their non–occurrence in Sāmoan.

2.3 Examples of ungrammatical noun incorporation in Sāmoan

2.3.1 Incorporation from a complement with a head distinct from the moved element

According to Baker, incorporation of a noun from a prepositional complement violates the ECP because the trace has a head distinct from the moved element. The following analysis shows why , in Baker's theory, movement from a prepositional complement violates the ECP.

25) *



In this analysis the first movement operation is blocked . The moved element c–commands its trace but there is a barrier to government since the head of the XP above B (PP) has a head (P) which is distinct from that of the moved element. Since this movement is blocked, incorporation cannot take place into V, making 26b) below ungrammatical.

26	a)	oloo	matamata	le	tama	i	le	televise
		PRES	watch		the boy	at/to	the	television
	*b)	oloo	matamata–televise		le	tama		
		PRES	watch–television		the	boy		
27	a)	sa	tao	le	moa	i	le	ogaumu
		PST	roast	the	chicken	in	the	oven
	*b)	sa	tao–oguamu		le	moa		
		PST	roast–oven		the	chicken		

28 a) sa latou malaga mai i le va'a
PST they travelled in the boat

*b) sa latou malaga–va'a mai
PST they travelled–boat

29 a) sa nofo le tama i le nofoe
PST sit the boy in the chair

*b) sa nofo–nofoa le tama
PST sit–chair the boy

30 a) sa taalo le tama i le paka
PST play the boy in the park

*b) sa taalo–paka le tama
PST play–park the boy

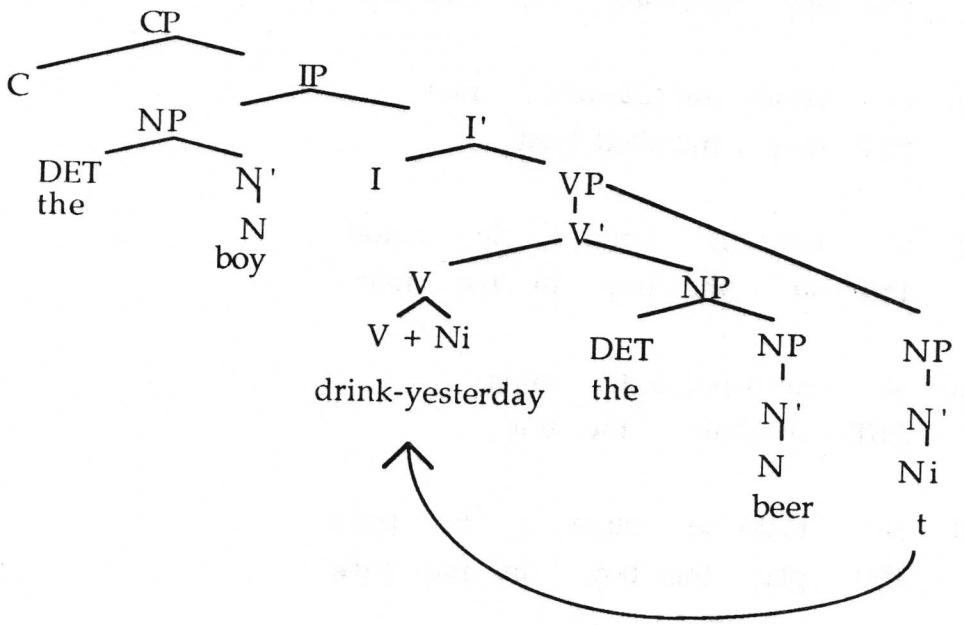
31 a) sa nofo le teine i le laulau
PST stand the girl on the table

*b) sa nofo–laulau le teine
PST stand–table the girl

2.3.2 Incorporation from a modifier

According to Baker, another reason for an ECP violation is when incorporation is from a modifier. Since the following all contain incorporations from time phrases, which are modifiers, they should all violate the ECP (see Baker 1988: 86).

32) *



33 a) na inu e le tama le pia ananafi
PST drank the boy the beer yesterday

*b) na inu–ananafi e le tama le pia
PST drank–yesterday the boy the beer

34 a) sa leoleo le tama i le po
PST guard the boy at the night

*b) sa leoleo–po le tama
PST night–watch the boy

PRES talk–boy the girl

35 a) E alu i le lotu i le Asosa
PRES he goes to church on Sunday

*b) E alu–assosa i le lotu
PRES he goes–Sunday to church

36 a) E faitautusi nanei
PRES he reads tonight

*b) E faitautusi–nanei
PRES he reads–tonight

37 a) sa matou malolo anapo
PST we rested last night

*b) sa matou–malolo anapo
PST we rested–last night

RESEARCH QUESTIONS AND HYPOTHESES

Reaction times, accuracy and certainty data on judgements about seven types of sentences were collected. The seven types, as previously discussed, relate to two distinct claims about the influence of language universals on knowledge of second language word formation rules. Types 1 to 3 are motivated by research questions concerning the influence of the implicational hierarchy proposed by Mithun (1984, 1986), and are summarized in Table 1 below. Type 1 is an unincorporated sentence. Type 2 corresponds to the first type of incorporation Mithun discusses, which is very productive in Sāmoan, detransitivizing noun compounding (see section 2.2.1). Type 3 corresponds to the most marked form of noun incorporation discussed by Mithun, classifier incorporation (see section 2.2.2). Types 4–7 are motivated by research questions concerning the influence of Universal Grammar and ECP violations and are

summarized in Table 2 below. Type 4 sentences are unincorporated sentences with complements, and Type 5 sentences are the counterpart ECP violations in which in which incorporation has taken place from a complement (see section 2.3.1). Type 6 sentences are unincorporated sentences with modifiers, and Type 7 sentences are the counterpart ECP violations in which incorporation has taken place from a modifier (see section 2.3.2).

a) Implicationally motivated sentence types

Some results from research on implicational universals, of the type identified for noun incorporation by Mithun, appear to support claims that more marked structures correlate with apparent areas of difficulty in second language acquisition (see e.g., Gass 1979; Pavesi 1986; Eckman, Bell & Nelson 1988; Doughty 1991 on the acquisition of relative clause formation and the influence of the noun phrase accessibility hierarchy). Eckman's markedness differential hypothesis (Eckman 1977: 1985) predicts that the more marked incorporation processes on Mithun's hierarchy will be more difficult to acquire for an L2 learner with an L1 with no productive examples of noun–verb incorporation, (as in English). Examples of classifier incorporation, the most marked form of noun incorporation identified by Mithun, were constructed in consultation with native Sāmoan speakers, (Type 3 in Table 1 below). Examples of the least marked form of incorporation identified by Mithun were also constructed, (Type 2 in Table 1 below). Finally, unincorporated forms of the Type 2 sentences were constructed (Type 1 in Table 1 below) giving three sentence types as a test of whether degree of difficulty suggested by the implicational relations was reflected in reaction time, certainty and accuracy measures.

Table 1

Implicationally motivated sentence types

TYPE 1	<	TYPE 2	<	TYPE 3
–incorporation		+incorporation		+incorporation
+occurring		+occurring		–occurring
–ECP violation		–ECP violation		–ECP violation
unincorporated		empty NP incorporation		classifier incorporation

b) UG motivated sentence types

There are contrary views of the influence of universal grammar on the development of second language syntactic knowledge. Bley–Vroman's fundamental difference hypothesis (1988) claims that access to initial UG parameters is cut off for adult second language learners, so claims about difficulty and stages of acquisition based on UG are of no relevance. On the other hand both Flynn (1991) and White (1989), amongst others, claim that there is evidence to support the continued availability of UG in second language acquisition, and therefore that difficulties and stages in L2 acquisition can be successfully predicted by UG theory. To test this hypothesis via evidence from reaction time, certainty and accuracy measures four sentence types were constructed. These are illustrated in Table 2 below.

Table 2

UG motivated sentence types

TYPE 4	TYPE 5	TYPE 6	TYPE 7
–incorporation	+incorporation (from a complement)	–incorporation	+incorporation (from a modifier)
+occurring –ECP violation complement structure	–occurring +ECP violation incorporated complement	+occurring –ECP violation modifier structure	–occurring +ECP violation incorporated modifier

c) Research Hypotheses for NNS responses

Research question 1: The effect of implicational relations on judgements.

Hypothesis 1: More marked examples i.e. Type 3 sentences, will elicit slower reaction times than Type 2 and Type 1 as an index of processing difficulty.

Hypothesis 2: Type 3 sentences will also elicit a greater number of uncertain responses than type 2 , and Type 2 will elicit a greater number than Type 1, since subjects are more likely to avoid categorial judgements in the case of difficult, marked items than they are where the items are less difficult and marked.

Hypothesis 3: More marked Type 3 sentences will also elicit a greater number of inaccurate responses than Type 2 sentences, with Type 1 sentences being most accurate.

Research Question 2: The effect of ECP violations on judgements

Hypothesis 4: The ECP violations, Types 5 and 7, will elicit slower reaction times than Types 4 and 6 as an index of processing difficulty.

Hypothesis 5: The ECP violations, Types 5 & 7 will elicit a greater number of uncertain responses, than Types 4 and 6 as an index of the difficulty of matching these sentences with possible structures in the language.

Hypothesis 6: Types 5 &7 will elicit a greater number of inaccurate responses than Types 4 & 6, in proportion to the degree to which learners are sensitive to the UG violations in those sentences, great inaccuracy being evidence for no sensitivity, marginal inaccuracy being evidence for sensitivity.

In all cases it is expected that NS responses will be faster, more certain and more accurate than NNS responses.

METHOD

Subjects: Subjects were 29 native speakers of English who had been enrolled for between two and three semesters in Sāmoan language programs at the University of Hawai'i at Manoa. A control group of 5 native Sāmoan speakers was used. The native speakers came from a variety of backgrounds. All had been brought up in Sāmoa, then had come to live in Hawai'i where they maintained their language through contact with other expatriate native speakers. The NS have had much less exposure to the written form of Sāmoan than the NNS students, who have had daily exposure over a period of at least two semesters through their instructional materials. All subjects reported themselves to be righthanded.

Materials: Reaction times, accuracy and certainty data were collected on responses to forty two sentences, randomized for each subject. There were six examples of each of the seven types described above (one example of type 6 and

7 sentences were wrongly classified and removed from the analyses). The sentence stimuli were constructed in consultation with two native speakers of Sāmoan, and the NNS subjects' teacher of Sāmoan to ensure both the validity of the grammatical/ ungrammatical distinctions, and to ensure that the NNS subjects would all be familiar with the lexical items used in each sentence. Lexical items used were, accordingly, all high frequency words that students had been exposed to in their Sāmoan language classes. Sentences were controlled for complexity and length to the extent that each consisted of a verb and two arguments[3]. The software used was 'Mindlab'. A MacIntosh computer & Qwerty keyboard were used for the data collection.

Procedures: Subjects were seated at a Macintosh SE computer and told that they were to see some sentences written in Sāmoan. These sentences would appear individually. They were told to respond to each sentence using one of three options, 'n' for not sure, 'm' for mistake or ungrammatical and 'c' for correct or grammatical. The convention for joining two words together into a single word using a hyphen was mentioned. Subjects were required to press the space bar three times in the course of the pretest instructions before the first stimuli appeared. The stimulus sentence remained on screen until the subject responded with a keystroke, whereafter it was replaced by the next sentence.

Measures taken: Reaction time and accuracy measures were taken of subject responses.In L1 sentence matching experiments reaction time has been claimed to correspond to degree of processing difficulty (Freedman & Forster 1985: Crain & Fodor 1987) where differences in reaction time have been explained as indices of the time taken to mentally parse compared sentences. Where one of the sentences is ungrammatical then normal parsing procedures are disrupted, resulting in a processing, and consequently a response delay. The reaction time data use in the present study, however, is more similar to that used by Cook (1990) than to that used by Freedman & Forster (1985) and Crain & Fodor (1987). Bley–Vroman & Masterson (1989) have argued for the use of reaction time measures as a supplement to grammaticality judgements in examining

[3]This is true of all the stimuli sentences with the exception, however, of examples 33a and 36b above which could be considered to consist of two arguments plus an adverbial modifier.

degree of processing difficulty in L2 rule application. Both Cook (1990) and Ellis (1991) have claimed that longer response times are indicative of greater degrees of difficulty in L2 grammaticality judgements.[4]

Arguments have been put forward for trimming the data in reaction time studies to avoid analyses based on skewed distributions (Heathcote, Pophiel & Mewhort 1991) though it is acknowledged that this can threaten the validity of conclusions drawn from the analyses. For this reason untrimmed data was analyzed in this study. Others have drawn attention to the fact that reaction time data can mask differences in individual or group strategies on certain tasks (Marquar & Pereira 1990) urging that it should be accompanied by other measures.

The other measure used in this study was a grammaticality judgement measure with a 'not sure' option which provided a measure of certainty. Grammaticality judgement measures have been traditionally used in studies of L2 competence (e.g., studies reported in White 1989) though doubts have been raised as to whether these reflect the operation of the computational mechanisms dedicated to language processing, or whether other influences such as general problem solving strategies (Ellis 1991) or judgements influenced by degree of previous exposure to the same or similar examples (Birdsong 1989; Robinson & Ha 1993; Robinson 1993; Nagata 1988) may be compounded with grammaticality judgements.

Response certainty options have been used in a number of previous L2 studies (see Chaudron 1983 and Ellis 1991 for a review). The present study uses them as a control for problems associated with grammaticality judgements

[4]Are errors faster, the same as, or slower than corresponding correct responses? Research on animal behaviour, using pigeons, shows that in quick, difficult visual discriminations errors are faster than correct responses– the larger the error the faster it is. For humans on visual discrimination tasks the opposite is true, errors are slower, and the slower the response the more likely it is to be an error. Swensson (1972) reports that for humans errors are faster than correct responses when two conditions are met– the discrimination is easy and the pressure to be fast is substantial. It is not clear, though, that this equation between speed and accuracy holds for second language grammaticality judgements, and it seems likely that would be influenced by such factors as linguistic markedness, degree of violation of grammaticality, and extent of first language distance from the stimuli, though this has yet to be demonstrated in SLA studies.

which require binary 'correct/ incorrect' responses. Sorace (1990) has observed that forcing this categorial distinction, ' may provide inaccurate or deceptive information about the learner's state of interlanguage competence, particularly if the object of investigation is an indeterminate structure '(Sorace 1990: 140, see also Schachter, Tyson & Diffley 1976). The 'not sure' option response leaves subjects free to express their uncertainty directly. If this option were not provided then it would be more difficult to claim that differences in response time were solely an index of processing difficulty, as is claimed in this study. In other words, one needs to avoid forcing learner's to hesitate in choosing between two unsatisfactory alternatives , correct/ incorrect, so that one can , with more conviction, claim that response delay is an index of the processing complexity of deciding on grammaticality where intuitions are clear and determinate.

RESULTS AND DISCUSSION

Implicational relations: Hypothesis 1, that Type 3 sentences would be slower than Type 2 is disconfirmed by the NNS and NS data, although both groups are slower on Type 3 sentences than Type 1 sentences (see Figure 1 and Figure 2). Hypothesis 2 is also disconfirmed. There are greater inaccuracies on Type 2 sentences than on Type 3 sentences for both groups, although both groups are most accurate on Type 1 sentences (see Figure 3). Hypothesis 3 is confirmed for the NNS group, responses to Type 3 sentences are much more uncertain than responses to Type 2 sentences (see Figure 4).

ECP violations: Hypothesis 4 is partly confirmed; there are significantly faster responses to Type 4 sentences than to Types 5 & 7 for both groups(see Figure 1 and Figure 2). However the NNS group respond only slowly to Type 6 sentences, while the NS group respond quickly to the ECP violations in Type 7 sentences. Hypothesis 5 is confirmed, both groups are more accurate in responses to Types 4 and 6 than they are to Types 5 and 7 (see Figure 3). Hypothesis 6 is confirmed for the non–native speakers, responses to Types 5 and 7 are much more uncertain than responses to Types 4 and 6 (see Figure 4). These results are discussed in more detail below.

Reaction time data

Means and standard deviations for the reaction times by sentence type for NNS and NS subjects are given in Figures 1 & 2 below. The variation in mean reaction time across sentence types appears similar for both groups. Type 3 sentences, classifier incorporation, take longer than the monotransitive unincorporated Type 1 sentences. However both groups take longest over Type 2, the occurring and grammatical examples of lexical compounding. This is strange for the NS group, particularly, since these are widely occurring in Sāmoan.

Birdsong, though, (1989) has noted (citing work on metalinguistic judgements of their L1 by native speaker illiterates, reported in Scribner & Cole (1982)) that naive native speakers tend to make grammaticality judgements based on acceptability, or previous experience of the occurrence of the stimuli in natural contexts, rejecting those stimuli which though grammatical seem unacceptable simply because never previously encountered. If it is this, and not grammaticality per se, that determines speed of responses in the case of some of the occurring and grammatical sentences, Types 1, 2, 4 & 5, it is still strange that responses to Type 2 sentences are slow for both groups. One would expect responses to Type 3 sentence to be slower, since they do not occur.

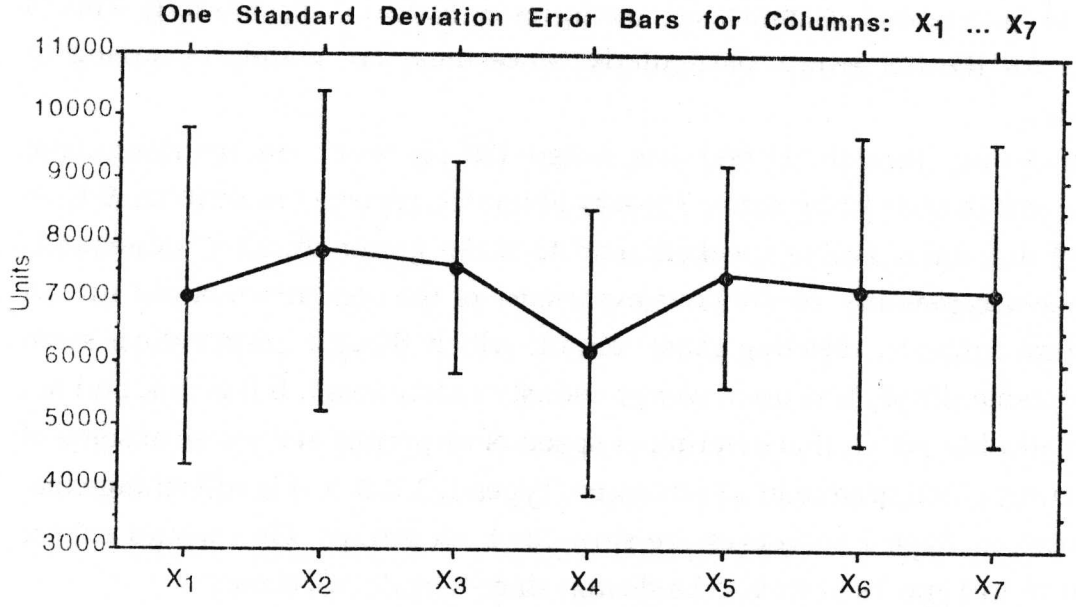
One would also expect responses to Type 3 sentences to be faster than responses to the other nonoccurring sentences since these are possible sentences, at least with respect to the constraint on possible and impossible forms discussed here, i.e. violation of the ECP. Type 3 sentences do not violate the ECP (that is, they are possible though nonoccurring forms) in contrast to Types 5 and 7 which are impossible and nonoccurring forms. Yet there is no significant difference in response time between Type 3, and Types 5 and 7 in either the NS or the NNS data. This implies a problem in using the response time data as an index of both acceptability, (defined as familiarity based on experience with similar occurring forms), and of grammatically defined possibility .

Both NNS and NS groups show the expected difference in reaction times between Type 4 and Type 5 sentences, with the ECP violations ,Type 5, eliciting

slower mean reaction times. However reaction times for both groups are identical with respect to Type 6 sentences and their counterparts, Type 7 sentences with ECP violations. NS responses to Types 4 and 6 have much lower standard deviations than responses to Types 5 and 7; this isn't true of NNS responses which have consistently wide standard deviations.

Figure 1

Mean reaction times and standard deviations for non–native speaker responses

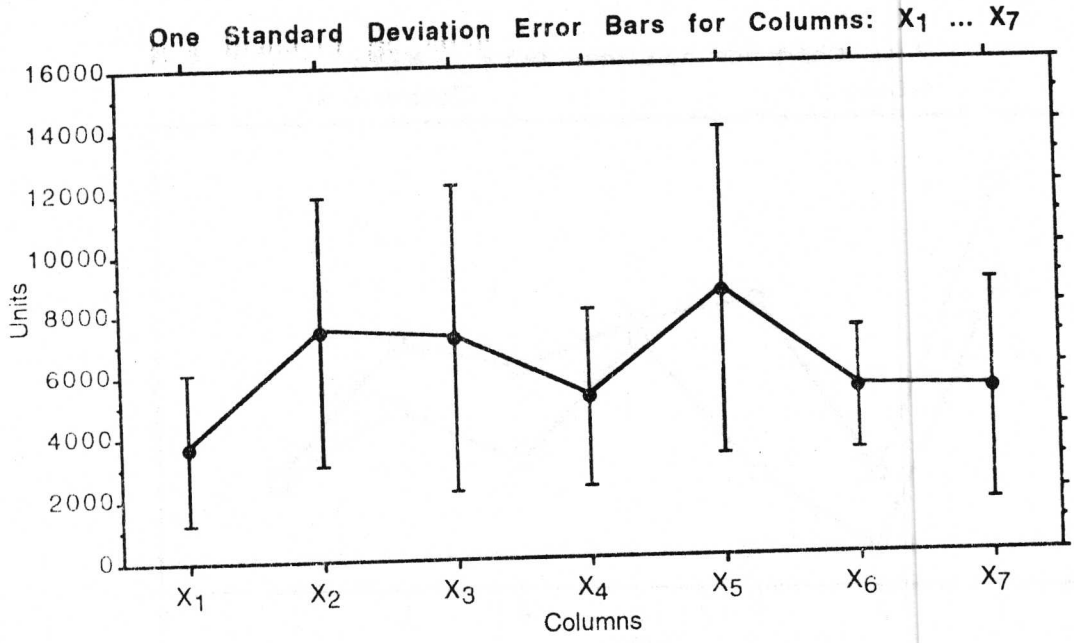


A one way ANOVA on mean reaction times to sentence types by NS /NNS groups shows significant between group differences only for reactions to Type 1 sentences (p= .0138, Scheffé 6.783) which are faster for the NS group than for the NNS group.

Separate one way repeated measures ANOVAs on reaction times to sentence types for NS and NNS groups show significant differences (p=.0001 in both cases). Post hoc comparisons show that for the NS group the significant differences are Type 1 v. Type 2 and 5, and Type 4 v. Type 5 . For the NNS group these differences are Type 4 v. all other types.

Figure 2

Mean reaction times and standard deviations for native speaker responses



Accuracy and certainty data

Percentage scores for accurate, inaccurate and not sure responses by sentence type were calculated for NS and NNS subjects . Figure 3 below gives percentage scores for accurate responses, and Figure 4 gives percentage scores for not sure responses.

Figure 3

Percentage scores for correct responses by native and non-native speakers

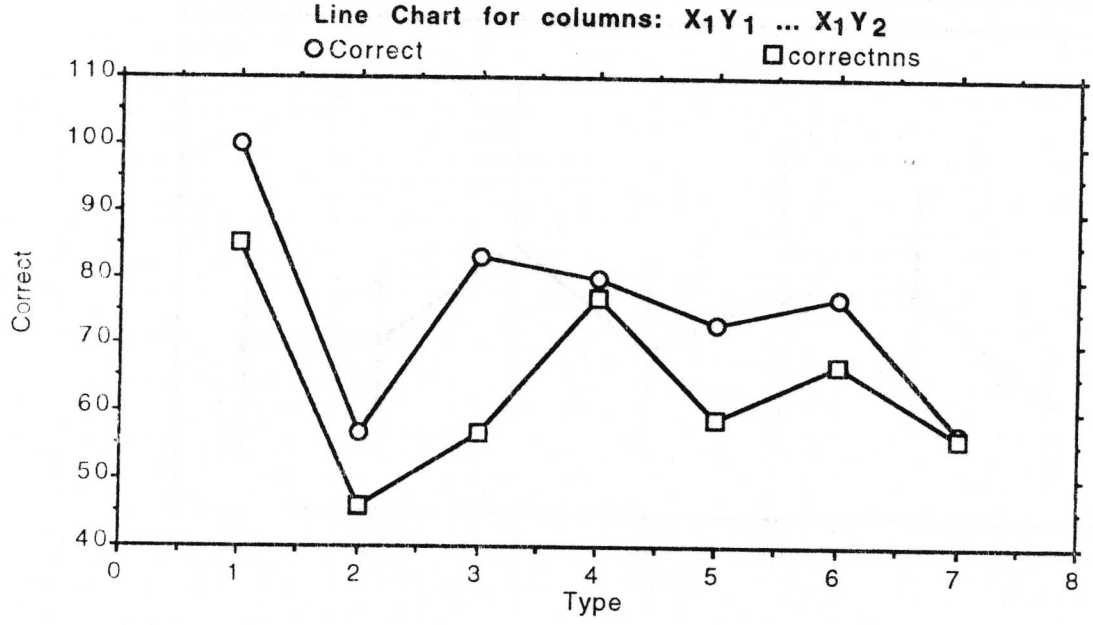
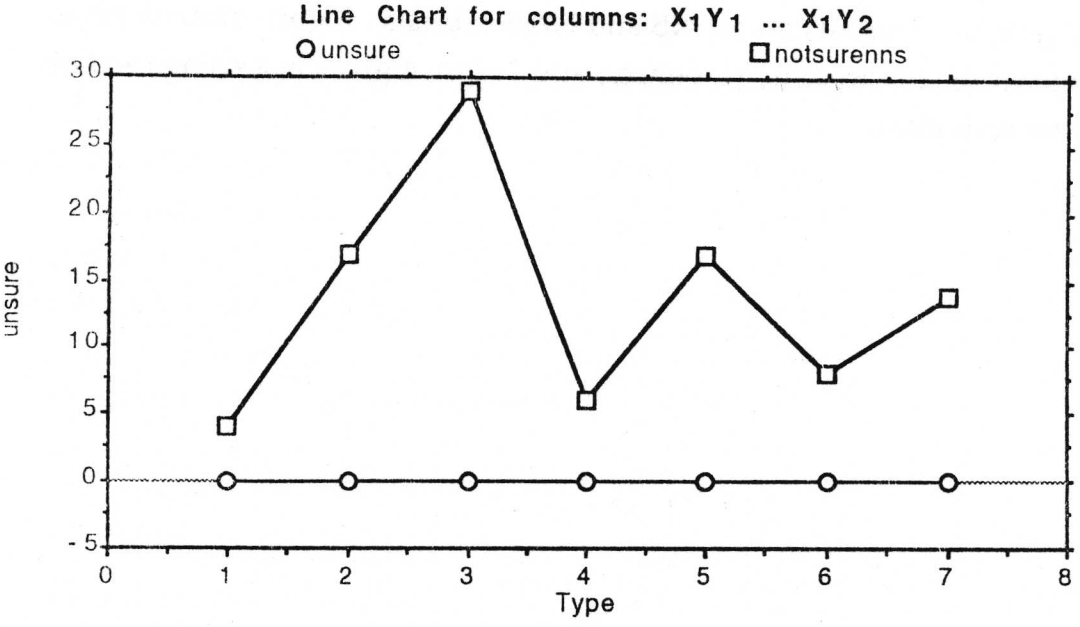


Figure 4

Percentage scores for not sure responses by native and non-native speakers



The certainty data in Figure 4, in contrast to the accuracy data in Figure 3, follows the pattern predicted by the research hypotheses. None of the NS subjects indicated uncertainty, whereas many of the NNS subjects did. The uncertainty responses for the NNS group follow the pattern predicted by the implicational hypothesis and the universal grammar hypothesis. There is an increase in uncertainty from type 1 to Type 3, following the implicational ordering suggested by Mithun (1984), while Type 5 and Type 7 elicit greater uncertainty than Types 4 and 6, as predicted by the ECP violations in 5 and 7.

The accuracy data, however, shows no difference between NS and NNS subjects in responses to Type 4 and 7 sentences, though there are clear differences between NS and NNS subjects on responses to Type 1, 2, 3 ,5 & 6 sentences, where NS subjects make more accurate responses. Accuracy is highest for both groups on Type 1 sentences and lowest on Type 2 sentences. This latter result is odd, for NS subjects particularly since, as previously stated, Type 2 sentences are widely occurring in Sāmoan. Like the NNS subjects NS subjects also tend to judge Type 7 sentences as grammatical, when they are not. However, since many of the modifiers in this set are temporal phrases, like 'yesterday' or 'at night' or 'on Sunday',and since Hodvhuagen (1985: 86) has observed that the position of temporal phrases in Sāmoan is relatively unfixed, it may be that native speakers consider this less of a violation than the incorporation from a complement in Type 5 sentences (for recent work on degrees of grammaticality in government and binding theory see Epstein 1990).

The accuracy data shows greater inaccuracy on ECP violations than non violations for both NS and NNS subjects, with NNS subjects making less accurate responses to these types than NS subjects (see Figure 3). Both groups tend to judge the violations acceptable. NNS subjects also show clear differences in the certainty of responses to Types 4 and 6, and Types 5 and 7, with the ECP violations eliciting more uncertainty than their counterpart sentences . Uncertainty judgements also clearly reflect the hypothesized order of difficulty for Types 1 to 3, as shown in the progressive uncertainty of responses from Type 1 to Type 3.

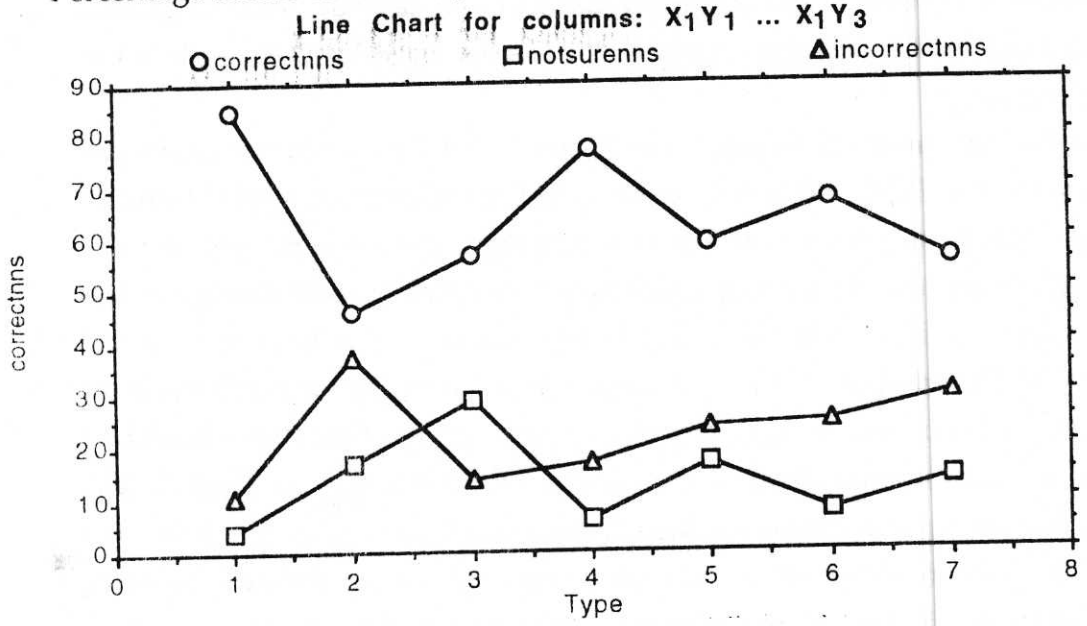
The relationship of certainty to accuracy data

To what extent does the NNS certainty data reflect the same difficulties in deciding on the grammaticality of the sentences that is reflected in the accuracy data? The relationship of the NNS certainty data to the accuracy data is shown in Figure 5 below. These results indicate a relationship between certainty and accuracy on responses to Types 1, 2 ,4, 5 & 7. The increase in uncertainty to responses to Type 2 relative to Type 1 sentences is accompanied by an increase in inaccuracy. This is in line with the claims of the implicational hypothesis relating degree of difficulty to degree of markedness. The increase in uncertainty to Type 5 and & 7 sentences, which are ECP violations, relative to Type 4 sentences, which are not, is similarly accompanied by an increase in inaccuracy of response to these items. This is in line with the prediction (RH6) that categorial judgements will be more difficult to make in the case of ECP violations.

These relationships between selection of the not sure option and inaccurate responses suggests that uncertainty is partly an artifact of the difficulty of making correct grammaticality judgements. However, responses to Type 3 sentences clearly run counter to this correlation between NNS uncertainty and inaccuracy since responses to these items are the most uncertain, though with the exception of responses to Type 1 sentences they are also the least inaccurate. Responses to Type 6 sentences could also be considered exceptions to this claim, since they elicit fewer not sure responses than Type 5 sentences, though inaccuracy on the two Types is at the same level. With these exceptions, though, the results suggest that for the NNS subjects uncertainty is related to the decisions they make about grammaticality, and that greater uncertainty is likely to reflect greater inaccuracy.

Figure 5

Percentage scores for correct, incorrect and not sure responses by NNS subjects



It is interesting to consider the relationship of accuracy to certainty measures in light of the distinction between possible and impossible nonoccurring forms drawn previously. In summarising the reaction time results it was pointed out that the slow reaction times for Type 3 sentences seemed odd when compared to Types 5 and 7, because, although all are nonoccurring in Sāmoan, Type 3 sentences are at least possible, since they violate no purported grammatical principles, such as the ECP. Type 3 sentences should therefore should obtain faster responses than Types 5 and 7, which are nonoccurring and impossible, but they do not.

One way of explaining the exceptional relationship between accuracy and certainty data for Type 3 sentences, is to invoke this distinction again. The relative accuracy of responses is accounted for by memory based knowledge that sentences of this Type do not occur, while the relative uncertainty of responses is the result of reconciling this knowledge with algorithmic knowledge that sentences of this type are certainly possible. In the cases of the other sentence types presented these two sources of information are not in conflict, i.e. what is possible does occur, and what is impossible does not (see Robinson & Ha 1993 for further discussion of the interaction between algorithm

driven processing and memory based processing of second language rules). In all other cases the two sources of information, and their indices (certainty data in the case of possibility, and accuracy data in the case of occurringness), confirm and so predict each other. Only in the case of Type 3 sentences does the information conflict.

Does the inaccuracy of responses to Types 5 and 7 mean that learners are not sensitive to the ECP, and is it then support for the fundamental difference hypothesis which suggests that second language learning progresses via general cognitive problem solving procedures mediated by knowledge of the first language? Possibly, but there is a complicating factor here not so far considered: if knowledge of the first language is an important influence on NNS decisions about correctness then it appears strange that they should be inaccurate on these items from a transfer perspective too. In English the paradigm case of lexical compounding, distinct from noun incorporation, is , for example 'book–reading' which allows compounds to be formed between verbs and objects of verbs but not verb plus subject as in 'librarian–reading', or complements as in 'yesterday–reading' (see Anderson 1988). So acceptance of Types 5 and 7 is odd from this perspective also , since the transfer explanation, like the UG explanation predicts that they would be reluctant to accept them as well–formed.

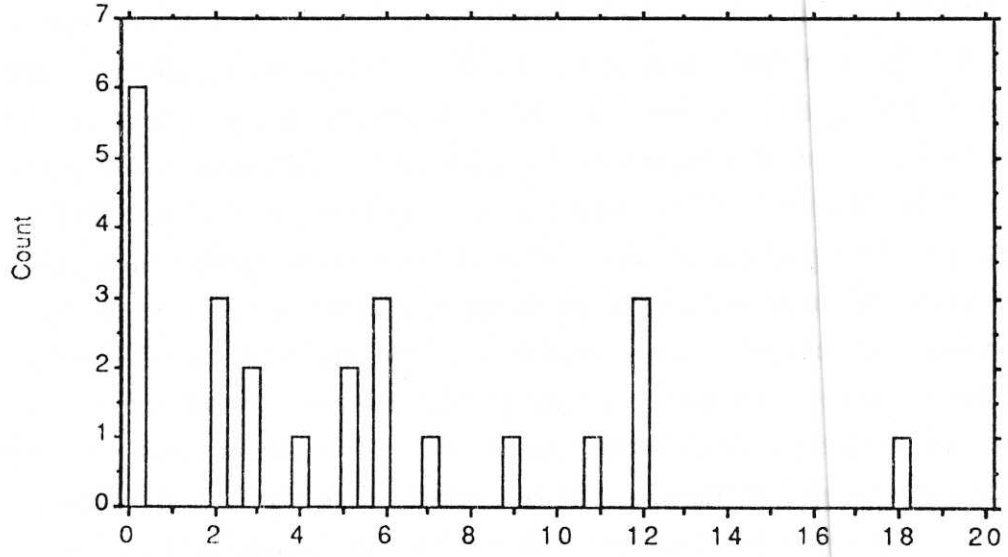
CONCLUSION

While the results of this study demonstrate once again (see Birdsong 1989; Bley–Vroman & Masterson 1989) the difficulty of making inferences about representations based on reaction time and grammaticality judgement data, they also suggest that certainty judgements are more sensitive to markedness and well-formedness than the other two measures. Although NS subjects are consistently faster and more accurate than NNS subjects across types of sentence,only decisions about response certainty clearly distinguish the native speaker and non native speaker groups. Certainty measures also support the research hypotheses: learners appear less willing to make categorial judgements about marked forms of word formation and those violating UG principles than they do about less marked forms and non UG violations. Sorace's objection to response certainty measures, that 'learners tend to fall into two major groups'

(1990:141)with regard to use of the not sure option– either using it all the time or never using it– was not the case here. Selection of the not sure option was distributed across NNS subjects responses in the following manner (see Figure 6 below), with only six subjects avoiding the not sure option, and ranging from two to eighteen selections for all other subjects.

Figure 6

Selection of the 'not sure' option by NNS subjects



In contrast to the earlier findings of Bley–Vroman, Ioup & Felix (1988), who reported minimal use of the 'not sure' option in their study, and of Ellis (1991), who reported that only 1.1 % of responses in a grammaticality judgement test of dative alternations made use of the not sure option, substantial numbers of not sure responses were recorded in this study, ranging from 3 % for Type 1 sentences to 30% for Type 3 sentences (see Figure 4).

Results of this measure, then, lend some support to the two claims which this study sought to address– that native English speaking learners would be sensitive to implicationally and UG motivated distinctions between types of word formation in Sāmoan. As acknowledged in the introduction to this paper, the extent of the influence of implicational universals, and of UG theoretic

syntactic constraints, on the acquisition of lexical knowledge in the L2 is an area where little research has so far been done. It is still predominantly syntactic and not lexical representation that SLA research has concerned itself with, despite the growing literature on this topic in mainstream psycholinguistics (see e.g., Marslen–Wilson 1989). The consequences of research in this area, though, for SLA theory and pedagogy are , potentially, at least as significant as research into the nature of syntactic representation.

Take, for example, two of the issues that follow from the present study. The first relates to the claims of the 'markedness generalization hypothesis' (see Eckman, Bell, Nelson 1988; Hamilton 1992). If lexical processes, like the variety of compounding examined here, are implicationally ordered, will teaching compounds lower on the hierarchy facilitate the acquisition of compounds higher up the hierarchy? The second issue relates to the possibility of incidentally inducing structural knowledge of a second language via a primary focus on its lexical meanings and word forms. If compounds are sensitive to the same restrictions governing the acceptability of syntactic structures, e.g., the ECP, might focussing on these sets of acceptable and unacceptable lexical items be a way of indirectly facilitating awareness of the larger configurational restrictions on syntactic structure? The relative difficulty of lexical processes in the L2,then, – for example the acquisition of word formation rules and the extent to which it is possible to induce syntactic knowledge of the L2 from evidence of its lexical structure – are areas of enquiry which are likely to be of great interest to SLA theorists and SL practitioners alike. It is to be hoped that future research, like that reported in the present study, will be directed towards answering these and other questions regarding the nature of L2 lexical processes, and to further motivating and refining the measurement instruments used in such research.

REFERENCES

Adams, V. (1973). *An introduction to modern English word formation*. London: Longmans.

Aitchison, J. (1987). *Words in the mind*. Oxford: Basil Blackwell.

Anderson, S. R. (1988). Typological distinctions in word formation. In Shopen (Ed.), *Language Typology and Syntactic Description* (Vol 3, ch 1 pp. 3–56).

Baker, M. (1988). *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Berwick, R. (1985). *The acquisition of syntactic knowledge*. Cambridge, Mass: MIT Press.

Birdsong, D. (1989). *Metalinguistic performance and interlinguistic competence*. Berlin, Heidelberg: Springer Verlag.

Bley–Vroman, R. (1989). The logical problem of foreign language learning. In S. Gass, J. Schachter (Eds), *Linguistic perspectives on second language acquisition*. New York: Cambridge University Press.

Bley–Vroman, R., Felix, S. & Ioup, G. (1988). The accessibility of Universal Grammar in adult second language learning. *Second Language Research 4*, 1–32.

Bley–Vroman, R. & D. Masterson (1989). Reaction time as a supplement to grammaticality judgement in the investigation of second language learner's competence. *University of Hawai'i at Manoa Working Papers in ESL, 7*, (2), 207–237.

Bley–Vroman, R. & C. Chaudron (1990). Second language processing of subordinate clauses and anaphora– First language and universal influences: A review of Flynn's research. *Language Learning, 40*, (2), 245–285.

Chaudron, C. (1983). Research on metalinguistic judgements: A review of theory, method, results. *Language Learning, 33*, 343–377.

Chomsky, N. (1982). *Lectures on government and binding*. Foris: Dordrecht.

Clark, E. V. (1981). Lexical innovations: How children learn to create new words. In W. Deutsch (Ed), *The child's construction of language*. London: Academic Press.

Clark, E. V. (1982). The young word maker. In E.Wanner, L. Gleitmen *Language acquisition: The state of the art*. New York: Cambridge University Press.

Clark, E. V. (1983). Meanings and concepts. In E. Markman (Ed.), *Handbook of child psychology Vol. III: Cognitive development*. New York: Wiley.

Clark E. V., & R. Berman (1984). Structure and use in word formation. *Language, 60*, (3), 542–589.

Cook, V. (1990). Timed comprehension of binding in advanced L2 learners of English. *Language Learning, 40*, 4: 557–599.

Crain, S. & Fodor, J. (1987). Sentence matching and overgeneration. *Cognition, 26*, 123–189.

Cutler, A. (1980). *Productivity in word formation*. Proceedings of the 16th regional meeting of the Chicago Linguistics Society.

DiScullio, A. & Williams, E. (1989). *On the Definition of Word*. Cambridge, Mass: MIT Press

Downing, P. (1977). On the creation and use of English compounds. *Language, 53*, (4), 810–842.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. Language Learning 24, (2): 315–330

Eckman, F., Bell, L., & Nelson, D. (1988). On the generalizability of relative clause instruction in the acquisition of English as a second language. *Applied Linguistics 5*, 138–155.

Ellis, R. (1991). Grammaticality judgements and second language acquisition. *Studies in Second Language Acquisition 13*, (2), 161–187.

Epstein, S. M. (1990). Differentiation and reduction in syntactic theory: A case study. *Natural Language and Linguistic Theory, 8*, 313–323.

Freedman, S. & Forster, K (1985). The psychological status of overgenerated sentences. *Cognition, 19*, 101–131

Flynn, S. (1987). *A parameter setting model of L2 acquisition: Experimental studies in anaphora*. Dordrect: Reidel.

Flynn, S. (1991). *Are second language learners in control*. Paper presented at the 12th Annual Second Language Research Forum, University of Southern California, March 1991.

Flynn ,S. & Lust, B. (1990). In defense of parameter setting in L2 acquisition: A reply to Bley–Vroman and Chaudron. *Language Learning, 40*, (3), 419–449.

Gass, S. (1979). Language transfer and universal grammatical relations. *Language Learning, 29,* (2), 327–344.

Givón, T. (1985). Universals of discourse structure and second language acquisition. In Rutherford, W (ed.), *Language universals and second language acquisition*. Amsterdam: John Benjamins.

Greenberg, J. H. (1966). *Universals of language* (2nd ed.). Cambridge, Mass: MIT Press.

Haiman, J. (1985) Natural Syntax Cambridge: Cambridge University Press

Hamilton, R. (1992). *The status of the markedness generalization hypothesis.* Unpublished manuscript, university of South Carolina Dept. Linguistics.

Heathcote, A., Popiel, S., & Mewhort, D. J. K. (1991). Analysis of response time distributions: An example using the stroop task. *Psychological Bulletin 109,* (2), 340–347.

Hovdhaugen, E. (1985). *A short sketch of Samoan grammar*. University of Helsinki.

Logan, G. (1988). Towards an instance theory of automatization. *Psychological Review, 95,* (4), 492–527.

Luce, R. D. (1986). *Response Times: Their role in inferring elementary mental organization*. Oxford: Oxford University Press.

Marchand, H. (1969). *The categories and types of present day English word formation*. Munich: Verlag C. H. Beck.

Marquar, J., & Pereira,M. (1990). Reaction Times in the Study of Strategies in Sentence–Picture Verification: A Reconsideration. *Quarterly Journal of Experimental Psychology, 42 A,* (1), 147–168.

Marslen–Wilson, W. (Ed.). (1989). Lexical Representation and Process Cambridge, Mass: MIT Press.

Mellow, D. (1989). *Noun incorporation in Cree*. Unpublished M. A. Dissertation, McGill University Department of Linguistics.

Mithun, M. (1984). The evolution of noun incorporation, *Language 60,* (4), 847–895.

Mithun ,M. (1986). On the nature of noun incorporation. *Language, 62,* (1), 32–38.

Nagata, H. (1988). The relativity of linguistic intuition: The effect of repetition on grammaticality judgements. *Journal of Psycholinguistic Research, 17,* (1), 1–28.

Olshtain, E. (1987) The acquisition of new word formation processes in second language acquisition. *Studies in Second Language Acquisition 9,* (2), 221–233.

Ouhalla, J. (1990). On incorporation. *Lingua 80,* (1), 55–77.

Romaine, S. (1983). On the productivity of word formation rules. *Australian Journal of Linguistics 3,* 177–200.

Robinson, P. J. (1993). *Reaction time data in second language acquisition research: a review of models and methods.* Unpublished manuscript, University of Hawai'i, Dept. ESL.

Robinson, P. J. & Ha, M. (1993). Instance theory and second language rule learning under explicit conditions. *Studies In Second Language Acquisition 15,* 4, 413–438.

Rosen, S. T. (1989). Two types of noun incorporation: A lexical analysis. *Language, 65,* (2), 294–317.

Sadock, J. (1986). Some notes on noun incorporation. *Language, 62,* (1), 19–31

Schachter, J., Tyson, A. F., & Diffley, F. J. (1976) Learner intuitions of grammaticality. *Language Learning, 26,* 67–76.

Shiffrin, W., & Schneider, R. M. (1977). Controlled and automatic human information processing: 1, Detection, search and attention. *Psychological Review, 84,* (1), 1–66.

Scribner, S & Cole, M. (1982). *The psychology of literacy.* Cambridge, MA: Harvard University Press.

Sproat, R. (1985). Infl and the configurationality of VSO languages. Manuscript.

Sorace, A. (1990). Indeterminacy in First and Second Languages: Theoretical and Methodological Issues. In J. de Jong & D. Stevenson (eds.), *Individualizing the assessment of language abilities.* Clevedon: Multilingual matters.

Swensson, T. (1977). Trade off bias and efficiency effects in serial choice reactions. *Experimental Psychology 95,* 397–407.

Talaa Samoa Anamua Part 2. (1975). *American Samoa Bicultural–Bilingual Project: Manuscript*, University of Hawai'i at Manoa, Dept. IndoPacific languages.

White, L. (1989). *Universal grammar and second language acquisition*. Oxford: Pergamon Press.

Williams, E. (1981). Language acquisition, markedness and phrase structure. In S. Tavoklian (Ed.), *Language acquisition and linguistic theory*. Cambridge, Mass: MIT Press.