Proceedings
of the Second Annual
High Desert Linguistics Society
Conference

Spring, 1999
University of New Mexico

Edited by Dawn Nordquist and Catie Berkenfield

Volume 2
1999

High Desert

Linguistics Society

TABLE OF CONTENTS

THE CONTINUITY OF "AGREEMENT": FROM PRE-LINGUISTIC ACTION GESTURES TO ASL VERBS[*]

SHANNON CASEY

University of California, San Diego

1. INTRODUCTION.

   Verb agreement morphology in American Sign Language (ASL), which is manifested through directional movement or spatial displacement of verb signs, is acquired relatively late (Meier 1982:109-12). The use of directional movement and spatial displacement is examined here in the gestural and sign productions of deaf children during their acquisition of ASL from their deaf parents. The findings support hypotheses for this late acquisition which pertain to verb-particular agreement differences (Lillo-Martin 1991, Newport and Meier 1985) and ASL's shifting spatial framework (Lillo-Martin 1991), but provide evidence against a hypothesis that the simultaneity of agreement hinders acquisition (Newport and Meier 1985).

   The term directionality is defined here as the use of movement, spatial displacement, and/or palm orientation in the production of a manual action gesture or sign to indicate an additional referent involved in an action. Many verbs in ASL can occur with one or more of three types of directionality.[1] In the first type, agreeing verbs indicate agents, patients, or recipients (Fischer and Gough 1980; Padden 1988, 1990).[2] In the second type, spatial

[1] There have been controversies concerning the linguistic status of pronouns and agreement in ASL (see Askins and Perlmutter 1995; Liddell 1995; Lillo-Martin 1986; Lillo-Martin and Klima 1990; Meier 1990; Padden 1988, 1990) and the distinction between agreeing and spatial verbs (see Fischer 1996; Janis 1992, 1995; Liddell 1990; Padden 1990; Shepard-Kegl 1985). These controversies and their implications for my analysis are addressed in my dissertation (Casey, forthcoming).
[2] Agreeing verbs are often described as agreeing with grammatical roles (e.g., subjects and objects). Semantic roles are used here (following the terminology of Goldin-Meadow and Mylander [1984] with the substitution of 'agent' for their 'actor') in describing the referent(s) of directionality (as proposed by Friedman [1975]) due to the fact that this study involves both sign and gestural productions. Many of the productions, especially of the youngest children, are nonlinguistic gestures or single ASL verbs which do not occur in sentences. Thus I do not wish to imply that these early productions are sentence-like through the use of terminology pertaining to grammatical roles. My use of semantic roles is meant to refer to pre-linguistic conceptions (as defined by Langacker for his role archetypes [1991:284-5]) and intended only to describe the role that the referent of directionality plays in the context in which the action gesture or sign is produced, and not to ascribe any awareness of semantic roles to the children.

verbs indicate locations (Padden 1988, 1990; Fischer 1996:104). While in the third type, some plain verbs can occur with pronoun clitics indicating agents, patients, or locations (Padden 1990). However, plain verbs involving body contact cannot occur with directionality (Padden 1990:122).

In ASL, if participants in an action are present in the immediate environment, directionality is produced with respect to the actual location of the referent. If participants are absent, places in space, called loci, are established for pronominal and directional reference to these participants (Friedman 1975:946, Klima and Bellugi 1979:276-7). For example, to say that Pat gives something to Kim, the verb GIVE moves from the locus of Pat (the agent) to the locus of Kim (the recipient) (see Figure 1).[3]

FIGURE 1. 'Pat gives Kim.' GIVE MOVES FROM THE LOCUS OF PAT ON THE LEFT TO THE LOCUS OF KIM ON THE RIGHT.
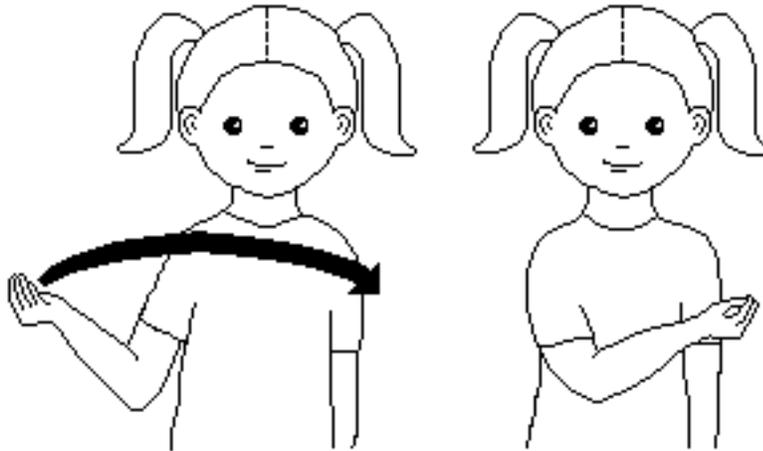
PAT-GIVE-KIM

---

[3] Following convention, ASL signs are represented by uppercase English glosses. If more than one English word is needed to gloss one sign, e.g., an ASL verb showing agreement, hyphens are used to join the English glosses.

2. PRIOR RESEARCH.

Petitto (1986) studied the acquisition of ASL pronouns, which resemble pointing gestures, in two deaf children of deaf parents. She found that these children went through the following four stages: 1) age 0;10-1;0—pointing at objects, locations, and people with a fully extended arm, i.e., extended outside of signing space; 2) age 1;0-1;6—pointing at objects, locations, and events, but not at people; 3) age 1;9-1;11—pointing to people keeping the arm within signing space, but with some reversal errors, e.g., YOU to mean ME; and 4) age 2;1-2;3—using YOU and ME pronouns correctly. This type of reversal error has also been found in the production of pronouns (Charney 1980; Chiat 1981, 1982) and agreement markers (Imedadze and Tuite 1992) during the acquisition of spoken languages. For example, a child may say "you want cookie" to mean "I want a cookie".

Previous study of the acquisition of verb agreement in ASL by deaf children of deaf parents has shown that it is not acquired with 90% correct production until age 3;0-3;6 (Meier 1982:109-12). This acquisition is relatively late in that it is acquired at about the same time as fusional morphology in spoken languages, whereas agglutinative morphology, as in Turkish, is acquired around age 2 (Newport and Meier 1985:929-30). Furthermore, it has been claimed that prior to the acquisition of agreement in ASL, children often produce verbs in an uninflected citation form (Hoffmeister 1978, Kantor 1982:99, Meier 1982:149).

On the other hand, people with no exposure to a signed language have been found to spontaneously produce action gestures containing movement to indicate additional referents, e.g., agents, patients, and locations. Deaf children (aged 1;4-7 yrs.) with no exposure to a signed system have been observed to produce action gestures containing directional movement and spatial displacement to indicate additional referents (Goldin-Meadow and Mylander 1990; Mohay 1982, 1990; Volterra et al. 1990). In another study, deaf children (approximately 10 yrs. old) exposed to Signing Exact English (SEE 2), which does not use directionality to indicate agreement, produced SEE 2 verb forms incorporating directional movement to indicate a verb's arguments (S. Supalla 1991). Hearing adults were found to produce directional action gestures to indicate participants in an action under experimental conditions in which they were asked to tell stories or describe videotaped scenes using gestures without speech (Casey and Kluender 1998; Dufour 1993). Hearing adults and children have also been shown to use directional action gestures simultaneously with speech (i.e., gesticulation) (Birdwhistell 1970:122-4, McNeill 1992). The production of directional gestures, e.g., an extended hand towards an object meaning 'give object', by hearing children acquiring spoken languages has also been observed (Bates et al. 1979, Blake and Dolgoy 1993, Petitto 1988—who also observed this in deaf children acquiring signed languages).

Based on these findings, I hypothesized that deaf children acquiring ASL may use directionality in their pre-linguistic action gestures and ASL verb signs to indicate additional referents to a greater extent than has previously been claimed. Additionally, children may go through stages similar to those found by Petitto (1986), in particular a stage in which they use directionality followed by a stage in which they produce only citation forms.

3. DIFFICULTY OF ACQUIRING VERB MORPHOLOGY IN ASL.

The late acquisition of ASL directional verb morphology has been hypothesized to be due to difficulties with the following:

a. with which verbs agreement can occur (Lillo-Martin 1991:162, Newport and Meier 1985:931-932), i.e., plain verbs with body contact cannot occur with agreement;

b. at which endpoint an argument should occur, i.e., some verbs move from the locus of the agent to the locus of the patient (HELP), whereas others move from the locus of the patient to the locus of the agent (HIRE) (Lillo-Martin 1991:162);

c. how many arguments with which a specific verb can agree, i.e., some can agree with two arguments (ASK), whereas others can only agree with one (SEE) (Lillo-Martin 1991:162);

d. with which semantic roles a verb can agree, i.e., some verbs occur with agent and patient agreement (LOOK-AT), whereas others occur with agent and recipient agreement (GIVE) (Lillo-Martin 1991:162);

e. shifting the spatial framework (Lillo-Martin 1991:162), e.g., a signer can take on the role of a third-person referent by shifting the shoulders toward that referent's locus and can subsequently use a first-person pronoun to indicate the referent whose role the signer has assumed (Lillo-Martin and Klima 1990:194-195);

f. simultaneity: In relation to Slobin's (1982) claim that morphology is easier to acquire if it is syllabic, stressed, and agglutinative, Meier (1982) and Newport and Meier (1985) hypothesized that the late acquisition of verb morphology in ASL is due to the *simultaneous* occurrence of this morphology with the verb stem, whereas morphemes which are acquired early in spoken languages are temporally and phonologically distinct from the stem (Meier 1982:151, Newport and Meier 1985:930).

Using data from deaf children's early productions, I will confirm the presence of directionality in their gestures and signs and analyze its use in terms of frequency of occurrence and error types. I argue that these data support hypotheses a.-e., but provide evidence against hypothesis f., i.e., that simultaneity can account for this late acquisition.[4]

---

[4] I am claiming that simultaneity is not problematic for the acquisition of directionality. However, Newport and Meier's (1985) hypothesis is also meant to account for the acquisition of path, manner, and combinations of central and secondary handshape classifiers in verbs of motion and location. I am not challenging the hypothesis that simultaneity is a factor in the late acquisition of these morphemes.

4. DATA.

The subjects were four deaf children (aged 0;8-2;11) of deaf parents acquiring ASL.[5] Productions of action gestures and signs from thirty-three videotaped free play and structured sessions (with the number of sessions for each child ranging from 6 to 14) were coded for the presence of directionality to indicate additional referents.[6] A reliability check between two coders, one of whom is a Deaf native signer, for a portion of the data indicated 97.68% agreement on whether or not an action gesture or sign was coded as containing directionality.

Whether a specific manual action production was coded as a gesture or sign was determined by:

      a. form—Did the production look like an ASL sign (taking into account that handshapes are the least accurate component of ASL signs produced by children aged 0;5-1;6 [Siedlecki and Bonvillian 1993]) or a commonly occurring gesture, e.g., open handed 'give' or raised arms 'pick (me) up'?

      b. contextual meaning—If it looked like an ASL sign, did that meaning fit the context? If it looked like a common gesture, did that meaning fit the context?

      c. child's age—If the child was under age 1;10, a production referring to an action was more likely to be a gesture, whereas if the child was over age 1;10, it was more likely to be a sign.

A reliability check between the two coders on a portion of the data showed 98.98% agreement on whether a production was a sign or gesture.

---

[5] An additional child's productions were analyzed, however his data were discarded because his input consisted of both ASL and Signing Exact English. His data are discussed in my dissertation (Casey forthcoming).

[6] The coding system used for these data was adapted mainly from the system I learned from Carolyn Mylander (see Goldin-Meadow et al. 1995), including features from Liddell and Johnson (1989), Meier (1982), and T. Supalla (1982).

5. RESULTS.

    Deaf children were found to produce action gestures and signs containing directionality to indicate additional referents at all ages studied. The children's spontaneous gesture and sign tokens, i.e., all of the action gestures and signs produced, were analyzed for the proportion of directional versus uninflected forms at each age (with the proportion for each child of a particular age averaged together). Overall, gestures were more often *directional* (see Figure 2), whereas signs were more often *uninflected* for additional referents at all ages (see Figure 3).

FIGURE 2. PROPORTIONS OF DIRECTIONAL VS. UNINFLECTED GESTURE TOKENS FOR ALL FOUR CHILDREN COMBINED. (Numbers above the bars show the actual number of gestures produced.)
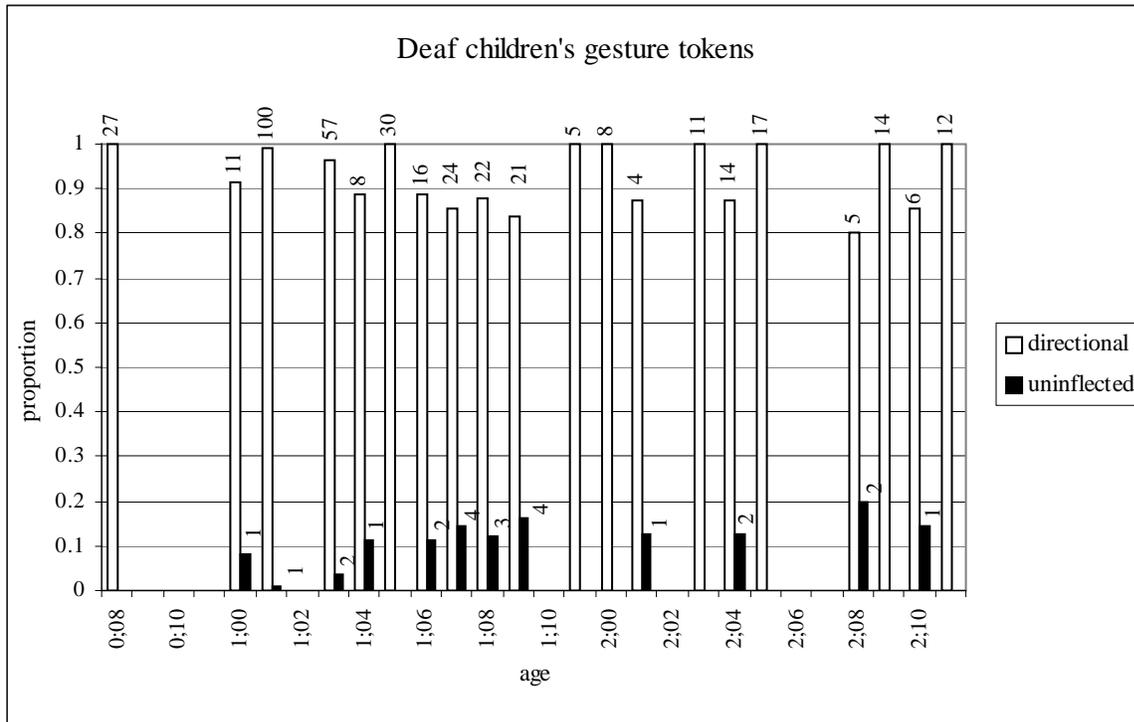
FIGURE 3. PROPORTIONS OF DIRECTIONAL VS. UNINFLECTED SIGN TOKENS FOR ALL
FOUR CHILDREN COMBINED. (Numbers above the bars show the actual number of signs
produced.)
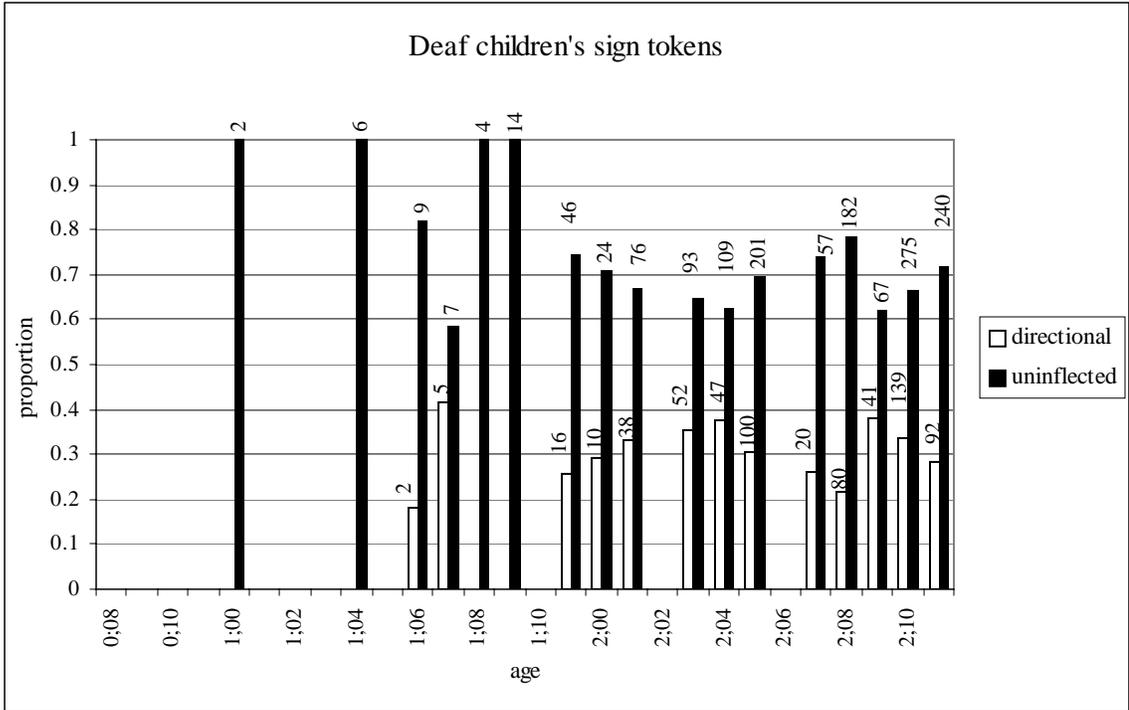
Similarly for types, i.e., only counting one occurrence of a particular ASL sign or a particular gesture meaning per session, gestures were more often *directional*, whereas signs were more often *uninflected*.

5.1. Obligatory Agreement. Some ASL verbs must occur with agreement. For example, the sign GIVE must occur with recipient agreement, while agent agreement is optional, however, the sign WANT may occur without agreement. The children's spontaneous productions were analyzed for the use of directionality only when it was obligatory and when the referents with which particular verbs must agree were present in the environment (following Meier 1982). Before the age of 1;11, there were only single tokens of verbs requiring agreement. However, from age 1;11-2;11 a greater percentage of verbs occurred with agreement (at least 70%) than without. (A paired t-test found this to be statistically significant: t(10) = 3.0329, *p* = .0126). The exception was age 2;7, at which only one verb requiring agreement was produced, and it was produced without agreement (see Fig. 4).[7]

---

[7] At age 2;4 more than 70% of verbs were produced with agreement, however, more tokens were produced without agreement (39), than with agreement (34). The reason for this apparent discrepancy is that this age combines the data from two children, one of whom produced 100% with agreement (3 tokens) and one of

FIGURE 4. PROPORTIONS OF VERB TOKENS WHICH OCCUR WITH VERSUS WITHOUT OBLIGATORY AGREEMENT FOR ALL FOUR CHILDREN COMBINED. (Numbers above the bars show the actual number of signs produced.)
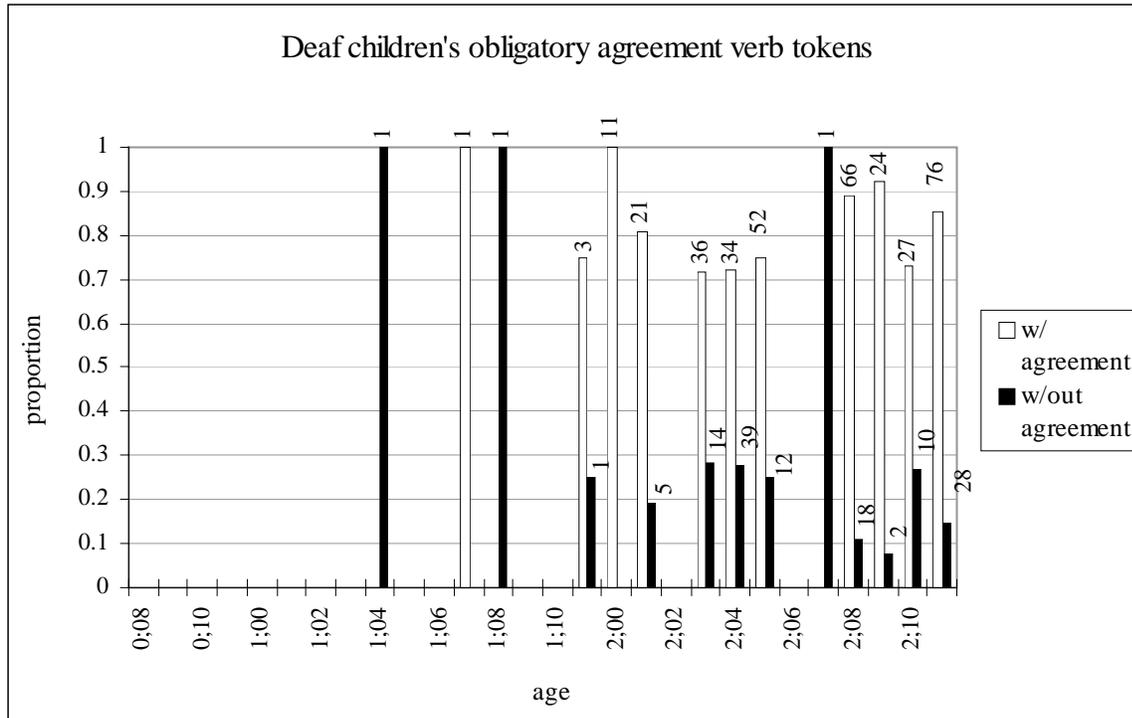
5.2. First Productions. ASL verbs (excluding plain verbs which cannot occur with a pronoun clitic) produced by a single child in more than one session were analyzed to determine if they were directional or uninflected during the first session in which they occurred. Table 1 shows the percentage of directional versus uninflected first productions with the number of types in parentheses. (One child did not produce any particular verbs occurring in more than one session):

whom produced 44.29% with agreement (31 tokens) and 55.71% without agreement (39 tokens). Thus, the average proportion of agreement produced at this age is 72.15%. This second child's lesser production of obligatory agreement is discussed in my dissertation (Casey, forthcoming).

TABLE 1. PERCENTAGES OF DIRECTIONAL VERSUS UNINFLECTED FIRST PRODUCTIONS.

	directional	*uninflected*
Maggie	52.63% (10)	47.37% (9)
Corinne	41.94% (26)	58.06% (36)
Ben	27.27% (9)	72.73% (24)

Maggie and Corinne did not show a great difference in the percentage of directional versus uninflected signs, but Ben produced more uninflected than directional.[8] However, a paired t-test (t(2) = -1.277, n.s.) of these data showed no statistically significant difference between the percentage of directional versus uninflected first productions.

5.3. Directionality errors. Children's production of directional errors was rare. Of the spontaneous verb signs produced by the four children only 4.9% contained directional errors (excluding errors of omission of obligatory agreement discussed in section 5.1). Most of these errors show difficulty with the first five areas hypothesized to be potential causes for the late acquisition of ASL:

       a. with which verbs agreement can occur (also see Bellugi 1988, Fischer 1973): Maggie (2;03) signed SLEEP produced toward a picture of sleeping dogs. SLEEP is a plain verb which cannot show agreement.

       b. at which endpoint an argument should occur (also see Meier 1982): To describe a scene in which a rabbit throws a stick at a pig, Ben (2;11) signed THROW moving toward a picture of a rabbit, but it should move from the rabbit toward the pig.

       c. how many arguments with which a specific verb can agree: Corinne (2;01) signed WANT moving from the direction of a tree toward herself to say that she wanted the tree. WANT can be displaced in space toward an agent or patient, but not cannot traverse space to show agreement with two arguments in the same production.

       d. with which semantic roles a verb can agree (also see Fischer 1973, Meier 1982): Corinne (2;01) signed FALL moving off of her chin to describe falling onto her chin. FALL can agree with a source and/or goal, but not with a body part onto which someone fell.

       e. shifting the spatial framework: In her description of a story in which a boy and a girl paint each other's faces and pour water over each

---

[8] Corinne is the same child in Meier (1982) and is the child named Carla in Petitto (1986).

other's heads, Maggie (2;11) ungrammatically used the opposite sides of her face to stand for different referents instead of using role shift, i.e., shifting her shoulders to take on the role of a third-person referent. For example, she used the right side of her face to describe the girl being painted by the boy, and the left side of her face to describe the boy being painted by the girl (see van Hoek et al. 1987 for discussion of this device used by other children).

The remaining errors resemble pronoun and agreement reversal errors. The data contain examples of children signing GIVE-YOU (the sign GIVE moving toward the addressee) to mean 'you give me' and ME-GIVE (the sign GIVE moving from the child) to mean 'you give me'. These errors were produced at age 1;11 by Maggie and ages 2;1 and 2;10 by Corinne. However, it is impossible to determine if these errors are comparable to pronoun reversal errors, i.e., the substitution of YOU to mean 'me', or if they involve problems with the endpoint at which an argument should occur (see b. above). Perhaps these children are not reversing the pronouns, but rather are incorrectly producing the recipient at the onset of the movement or the agent at the termination. On the other hand, there is one example which cannot involve endpoint confusions (as in b. above) and seems to clearly indicate a reversal, because the entity with which the verb agrees is not involved in the action. At 2;1 Corinne signed GIVE-ME (the sign GIVE moving toward herself) to mean 'give sister' in a context meaning 'grandmother give sister', i.e., she used agreement with herself meaning 'sister/her'. This example is produced at an age close to those at which Petitto (1986) found pronoun reversal errors in two children, (ages 1;9-1;11), one of whom was Corinne.[9]

6. DISCUSSION.

In contrast to my initial hypothesis, a stage in which children produce only citation forms of signs was not found. Thus, rather than the discontinuity between pointing gestures and ASL pronouns found by Petitto (1986), the uninterrupted use of directionality indicates a continuity between gestures and signs. Based on the above findings, I hypothesize that the late acquisition of directional verb morphology is not due to *simultaneity* for the following reasons:

a. *Simultaneous* directionality is used in gestures and signs throughout acquisition and for the same referential function as directional verb morphology in ASL.

b. The *majority* of ASL verb tokens which take obligatory agreement are produced with that agreement (see Lillo-Martin et al. [1998] for data in which even fewer omissions of obligatory agreement were found).

c. First productions of particular ASL verbs (excluding verbs which cannot occur with directionality) are *not more likely to be uninflected*

---

[9] Thanks to Karen Emmorey for pointing this out to me.

than directional. If simultaneity were difficult, children would be expected to produce uninflected forms of particular verbs before directional forms.

d. Errors can be accounted for by difficulties which are distinct from simultaneity, i.e., by the other hypotheses outlined above and by reversal errors.

Contrary to Petitto's (1992:55) assertion that "aspects of the structural and conceptual underpinnings of children's knowledge and use of language are fundamentally distinct from their knowledge and use of gesture", the continuous use of directionality for referential purposes throughout acquisition suggests that this use of directionality has gestural origins, although it has become grammaticized in ASL and other signed languages (see Armstrong et al. [1995] for a similar proposal concerning syntax). The hypothesis that the referential use of directionality has a gestural origin (Casey and Kluender 1995) can account for its occurrence in: 1) practically all signed languages studied so far (T. Supalla and Webb 1995); 2) the emergence of a new signed language in Nicaragua (Kegl and Iwata 1989, Senghas 1995); 3) gestures of deaf children with no ASL input (Goldin-Meadow and Mylander 1990; Mohay 1982, 1990; Volterra et al. 1990; S. Supalla 1991); and 4) gestures of hearing adults and children (Bates et al. 1979, Birdwhistell 1970, Blake and Dolgoy 1993, Casey and Kluender 1998, Dufour 1993, McNeill1992, Petitto 1988).

REFERENCES

ARMSTRONG, DAVID F., WILLIAM C. STOKOE, and SHERMAN E. WILCOX. 1995. *Gesture and the Nature of Language*. Cambridge: Cambridge University Press.

ASKINS, D. M., and D. M. PERLMUTTER. 1995. Allomorphy explained through phonological representation: Person and number inflection of American Sign Language verbs. Manuscript, University of California, San Diego and University of Rochester.

BATES, ELIZABETH, LAURA BENIGNI, INGE BRETHERTON, LUIGIA CAMAIONI, and VIRGINIA VOLTERRA. 1979. *The Emergence of Symbols: Cognition and Communication in Infancy*. New York: Academic Press.

BELLUGI, URSULA. 1988. The Acquisition of a Spatial Language. *The Development of Language and Language Researchers: Essays in Honor of Roger Brown*, ed. by Frank S. Kessel, 153-85. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

BIRDWHISTELL, RAY L. 1970. *Kinesics and Context: Essays on Body Motion Communication.* Philadelphia: University of Pennsylvania Press.

BLAKE, JOANNA, and SUSAN J. DOLGOY. 1993. Gestural Development and its Relation to Cognition During the Transition to Language. *Journal of Nonverbal Behavior* 17(2).87-102.

CASEY, SHANNON. Forthcoming. *"Agreement" in Gestures and Signed Languages: The Use of Directionality to Indicate Referents Involved in Actions*. Ph.D. Dissertation, University of California, San Diego.

CASEY, SHANNON, and ROBERT KLUENDER. 1995. Evidence for Intermediate Forms in the Evolution of Language. *Papers from the 31st Regional Meeting of the Chicago Linguistic*

Society, vol. 1: The Main Session, ed. by Audra Dainora, Rachel Hemphill, Barbara Luka, Barbara Need, and Sheri Pargman, 66-80. Chicago: Chicago Linguistic Society.

CASEY, SHANNON, and ROBERT KLUENDER. 1998. Pre-linguistic features of gesture in hearing adults. Poster presented at the Sixth International Conference on Theoretical Issues in Sign Language Research, Gallaudet University, November 14, 1998.

CHARNEY, ROSALIND. 1980. Speech Roles and the Development of Personal Pronouns. *Journal of Child Language* 7 (3).509-28.

CHIAT, SHULAMUTH. 1981. Context-specificity and Generalization in the Acquisition of Pronominal Distinctions. *Journal of Child Language* 8 (1).75-91.

CHIAT, SHULAMUTH. 1982. If I Were You and You Were Me: The Analysis of Pronouns in a Pronoun-reversing Child. *Journal of Child Language* 9 (2).359-79.

DUFOUR, ROBERT. 1993. *The Use of Gestures for Communication Purposes: Can Gestures Become Grammatical?* Ph.D. Dissertation, University of Illinois, Urbana-Champaign.

FISCHER, SUSAN D. 1973. Verb inflections in American Sign Language and their acquisition by the deaf child. Paper presented at the Linguistic Society of America Annual Meeting, San Diego.

FISCHER, SUSAN D. 1996. The Role of Agreement and Auxiliaries in Sign Language. *Lingua* 98.103-19.

FISCHER, SUSAN D., and BONNIE GOUGH. 1980. Verbs in American Sign Language. *Sign and Culture: A Reader for Students of American Sign Language*, ed. by William Stokoe, 149-79. Silver Spring, Maryland: Linstok Press.

FRIEDMAN, LYNN A. 1975. Space, Time, and Person Reference in American Sign Language. *Language* 51 (4).940-61.

GOLDIN-MEADOW, SUSAN, and CAROLYN MYLANDER. 1984. Gestural Communication in Deaf Children: The Effects and Noneffects of Parental Input on Early Language Development. *Monographs of the Society for Research in Child Development* 49(3-4).1-151.

GOLDIN-MEADOW, SUSAN, and CAROLYN MYLANDER. 1990. Beyond the Input Given: The Child's Role in the Acquisition of Language. *Language* 66.323-55.

GOLDIN-MEADOW, SUSAN, CAROLYN MYLANDER, and CYNTHIA BUTCHER. 1995. The Resilience of Combinatorial Structure at the Word Level: Morphology in Self-styled Gesture Systems. *Cognition* 56.195-262.

HOFFMEISTER, ROBERT JAMES. 1978. *The Development of Demonstrative Pronouns, Locatives and Personal Pronouns in the Acquisition of American Sign Language by Deaf Children of Deaf Parents*. Ph.D. Dissertation, University of Minnesota.

IMEDADZE, NATELA, and KEVIN TUITE. 1992. The Acquisition of Georgian. *The Crosslinguistic Study of Language Acquisition, Vol. 3*, ed. by Dan Isaac Slobin, pp. 39-109. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

KANTOR, REBECCA. 1982. *Communicative Interaction in American Sign Language Between Deaf Mothers and Their Deaf Children: A Psycholinguistic Analysis*. Ph.D. Dissertation, Boston University.

KEGL, JUDY, and GAYLA A. IWATA. 1989. Lenguaje de Signos Nicaragüense: A Pidgin Sheds Light on the "Creole" ASL. *Proceedings of the Fourth Annual Meeting of the Pacific Linguistics Conference*, 266-94. University of Oregon, Eugene.

KLIMA, EDWARD S., and URSULA BELLUGI. 1979. *The Signs of Language*. Cambridge, MA: Harvard University Press.

LANGACKER, RONALD W. 1991. *Foundations of Cognitive Grammar, Volume 2: Descriptive Application*. Stanford, CA: Stanford University Press.

LIDDELL, SCOTT K. 1990. Four Functions of a Locus: Reexamining the Structure of Space in ASL. *Sign Language Research: Theoretical Issues*, ed. by Ceil Lucas, pp. 176-98. Washington, D.C.: Gallaudet University Press.

LIDDELL, SCOTT K. 1995. Real, Surrogate, and Token Space: Grammatical Consequences in ASL. In *Language, Gesture, and Space*, ed. by Karen Emmorey and Judy Reilly, pp. 19-41. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

LIDDELL, SCOTT K., and JOHNSON, R. E. 1989. American Sign Language: The Phonological Base. *Sign Language Studies* 64.195-277.

LILLO-MARTIN, DIANE. 1986. *Parameter Setting: Evidence from Use, Acquisition, and Breakdown in American Sign Language*. Ph.D. Dissertation, University of California, San Diego.

LILLO-MARTIN, DIANE. 1991. *Universal Grammar and American Sign Language: Setting the Null Argument Parameters*. Dordrecht: Kluwer Academic Publishers.

LILLO-MARTIN, DIANE, RONICE DE QUADROS, and GAURAV MATHUR. 1998. Acquisition of verb agreement in ASL and LIBRAS: A cross-linguistic study. Poster presented at the Sixth Theoretical Issues in Sign Language Research, Gallaudet University, November 13, 1998.

LILLO-MARTIN, DIANE, and EDWARD S. KLIMA. 1990. Pointing out Differences: ASL Pronouns in Syntactic Theory. In *Theoretical Issues in Sign Language Research, Vol. 1: Linguistics*, ed. by Susan D. Fischer and Patricia Siple, pp. 191-210. Chicago: University of Chicago Press.

MCNEILL, DAVID. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.

MEIER, RICHARD P. 1982. Icons, Analogues, and Morphemes: *The Acquisition of Verb Agreement in American Sign Language*. Ph.D. Dissertation, University of California, San Diego.

MEIER, RICHARD P. 1990. Person Deixis in American Sign Language. *Theoretical Issues in Sign Language Research, Volume 1: Linguistics*, ed. by Susan D. Fischer and Patricia Siple, pp. 175-90. Chicago: University of Chicago Press.

MOHAY, HEATHER. 1982. A Preliminary Description of the Communication Systems Evolved by Two Deaf Children in the Absence of a Sign Language Model. *Sign Language Studies* 34.73-90.

MOHAY, HEATHER. 1990. The Interaction of Gesture and Speech in the Language Development of Two Profoundly Deaf Children. *From Gesture to Language in Hearing and Deaf Children*, ed. by V. Volterra and C. J. Erting, pp. 187-204. Berlin: Springer-Verlag.

NEWPORT, ELISSA, and RICHARD P. MEIER. 1985. The Acquisition of American Sign Language. *The Crosslinguistic Study of Language Acquisition, Vol. 1: The Data*, ed. by Dan Isaac Slobin, pp. 881-938. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

PADDEN, CAROL A. 1988. *Interaction of Morphology and Syntax in American Sign Language*. New York: Garland Publishing, Inc.

PADDEN, CAROL A. 1990. The Relation Between Space and Grammar in ASL Verb Morphology. *Sign Language Research: Theoretical Issues*, ed. by Ceil Lucas, pp. 118-32. Washington, D.C.: Gallaudet University Press.

PETITTO, LAURA ANN. 1986. *From Gesture to Symbol: The Relationship Between Form and Meaning in the Acquisition of Personal Pronouns in American Sign Language*. Bloomington, Indiana: Indiana University Linguistics Club.

PETITTO, LAURA ANN. 1988. "Language" in the Prelinguistic Child. *The Development of Language and Language Researchers: Essays in Honor of Roger Brown*, ed. by Frank S. Kessel, pp. 187-221. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

PETITTO, LAURA ANN. 1992. Modularity and Constraints in Early Lexical Acquisition: Evidence from Children's Early Language and Gesture. *Modularity and Constraints in Language and Cognition: The Minnesota Symposium on Child Psychology*, ed. by Megan R. Gunnar and Michael Maratsos, pp. 25-58. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

SENGHAS, ANN. 1995. *Children's Contribution to the Birth of Nicaraguan Sign Language*. Ph.D. Dissertation, Massachusetts Institute of Technology.

SIEDLECKI, THEODORE JR., and JOHN D. BONVILLIAN. 1993. Location, Handshape, and Movement: Young Children's Acquisition of the Formational Aspects of American Sign Language. *Sign Language Studies* 78.31-52.

SLOBIN, DAN I. 1982. Universal and Particular in the Acquisition of Language. *Language Acquisition: The State of the Art*, ed. by Eric Wanner and Lila R. Gleitman, pp. 128-70. Cambridge: Cambridge University Press.

SUPALLA, SAMUEL JAMES. 1991. Manually Coded English: The Modality Question in Signed Language Development. *Theoretical Issues in Sign Language Research, Volume 2: Psychology*, ed. by Siple and Fischer, pp. 85-109. Chicago: University of Chicago Press.

SUPALLA, TED. 1982. *Structure and Acquisition of Verbs of Motion and Location in American Sign Language*. Ph.D. Dissertation, University of California, San Diego.

SUPALLA, TED, and REBECCA WEBB. 1995. The Grammar of International Sign: A New Look at Pidgin Languages. *Language, Gesture, and Space*, ed. by Karen Emmorey and Judy Reilly, pp. 333-52. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

VAN HOEK, KAREN, LUCINDA O'GRADY, and URSULA BELLUGI. 1987. Morphological Innovation in the Acquisition of American Sign Language. *Papers and Reports on Child Language Development, Vol. 26*, pp. 116-23. Stanford, CA: Leland Stanford Junior University.

VOLTERRA, V., S. BERONESI, and P. MASSONI. 1990. How Does Gestural Communication Become Language? *From Gesture to Language in Hearing and Deaf Children*, ed. by V. Volterra and C. J. Erting, pp. 205-16. Berlin: Springer-Verlag.

SYNCHRONIC AND DIACHRONIC PERSPECTIVES ON NEGATIVE MODALS IN ASL

BARBARA SHAFFER

University of New Mexico

1. INTRODUCTION.

As has previously been shown (Shaffer 1998, Wilcox and Wilcox 1995), markers of modality in ASL are primarily free grammatical morphemes. The modal auxiliaries most often noted for ASL are: CAN (POSSIBLE), MUST (SHOULD), FUTURE, SEEM, FEEL, and OBVIOUS. There are in ASL, as there are in other documented languages, ways to express the negation of the modal notions. This paper will address those negative modal forms and, after analyzing one such form, offer some diachronic and synchronic conclusions regarding the development of negative modal notions in ASL.

The hypothesis is that markers of modality in ASL developed along predictable grammaticization paths described for modality in other languages. For example, in a discussion of markers of possibility, Bybee et al. (1994:190) note that there are several known cases of auxiliaries predicating physical ability that come to be used to mark general ability as well. Two cases are cited. English *may* was formerly used to indicate physical ability and later came to express general ability. The second case noted is Latin *potere* or *possum* 'to be able' which is related to the adjective *potens* meaning 'strong or powerful', and which develops into French *pouvoir* and Spanish *poder*, both meaning 'to be able' (1994:190). Wilcox and Wilcox (1995) and Shaffer (2000) have suggested that a similar grammaticization path can be seen for markers of possibility in ASL as well. Evidence from 1855 LSF, from which ASL developed (and which will be referred to as 1855 ASL in this paper), suggests that the lexical sign STRONG has grammaticized into the sign CAN which is used to indicate physical ability, mental ability, and root possibility, as well as permission and epistemic possibility (*Cf.* Brouland 1855 for the source data).

Long (1918) describes the articulation of POSSIBLE (CAN) and STRONG below:

> CAN—Expressing possibility, power, etc.—Hold the "S" hands out in front, elbows against sides, and let the hands drop a little way with a jerk. (1918:25)

> POWER, STRENGTH—Hold out the fists in front from the sides, elbows at sides; lift the fists toward the right, and throw them over the left with a circular or swinging motion and bring them down; keep them in the same relative position as to distance and direction of motion all the time. (1918:90)

> STRONG—Hold out the "S" hands to the front from the side; moving them slightly to one side, describe a small arc or circle, making a show of

using considerable force. The motion is something like slinging a sledge hammer. (1918:109)

Evidence from existing 1855 ASL sources supports the claim that CAN is a grammaticization of STRONG. In fact, the 1913 data would appear irrefutable.[1] McGregor, in a 1913 lay sermon (in SMI 1997), signs the following:

(1) NOW EACH OTHER BETTER AND WE CAN UNDERSTAND
    EACH OTHER BETTER AND FEEL BROTHER
       "We know each other better and are able to understand each other
       better and feel like brothers."

(2) OUR FATHER STRONG OVER MOON STARS WORLD
    "Our father has power over the moon, and stars, and world."

(3) SELF CAN GET ALONG WITHOUT OUR HELP
    "He can get along without our help."
            or
    "He is powerful without our help."

In (3) it is unclear whether the signer was intending a strength or possibility reading. Either meaning is possible and logical in sentence (3). This ambiguity provides good evidence for the development of CAN in ASL out of STRONG. In each of the above articulations the sign STRONG and the sign CAN are signed in an identical manner as shown in Figure 1 below. Further, in an 1856 dictionary the sign below was glossed both POUVOIR and COURAGE (Pélissier 1856).

FIGURE 1. OLD LSF CAN AND STRONG

---

[1] Most of the diachronic data was taken from The Preservation of American Sign Language—The Complete Historical Collection, SMI 1997.

In Shaffer (2000) I suggest the following path for modern ASL CAN:

FIGURE 2. THE SEMANTIC DEVELOPMENT OF ASL CAN

physical ability $\longrightarrow$ non-physical ability $\longrightarrow$ root possibility $\longrightarrow$ epistemic
 mental ability possibility

permission
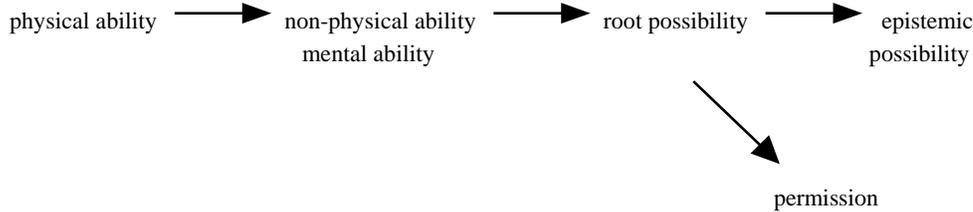
2. THE SEMANTICS OF CAN'T IN ASL.

While grammaticization paths can be hypothesized for other ASL modals as well, it is striking that the negation of modals usually results in a completely different form as well as a differing grammaticization path. In the case of CAN, for example, it might be assumed that, to indicate inability or to indicate refusal of permission, a simple negation of CAN resulting in CAN+NOT or NOT+CAN might be seen. This is not the case. In fact, there is no evidence, either diachronic or synchronic, of the use of CAN+*negative* seen for ASL. ASL data collected from the turn of the century as well as from dictionaries from LSF in 1855 to ASL in 1918 indicate that the means of expressing CAN+*negative* was the form commonly glossed as CAN'T, which is still in use today. There is written evidence of this form dating back to 1855 in France. This form is described by Long (1918) below and is seen as signed by Vedetiz (1913 [in SMI 1997]) in (4) and McGregor (1913 [in SMI 1997]) in (5) and (6):

> CAN'T—Holding the left "G" hand out in front, strike the end of it with the forefinger of the right "G" hand, as if cutting it off, and letting the right hand continue down. (1918:25)

> (4) 3rd PL NOT UNDERSTAND SIGN
>   BECAUSE PRO.3 PL CAN'T++ SIGN PRO.3 PL TELL
>   ANNOUNCE HERE SIGN REMOVE
>   "They do not understand sign and because they can't sign they declare sign banished."

> (5) [DURING THAT TIME]-top WORLD SEEM (2h) BIG PEOPLE
>   FEEL FAR FAR FAR FEEL INTERACT CAN'T CAN'T CAN'T
>   "At that time the world seemed large. People felt separate from each other, and felt that they could not interact."

(6) CAN'T FEEL BROTHER UNDERSTAND NOTHING JESUS HIS
WORD
"They cannot feel like brothers and do not understand the word of
Jesus."

As can be seen, by 1913 the above CAN'T already exhibited a range of meaning. In the first example (4) CAN'T is interpreted as a mental inability. In essence the signer is indicating that the people in question don't know how to sign. In (5) and (6) the meaning of CAN'T is interpreted as a negative root possibility, which I suggest is a further grammaticization from the original meaning of CAN'T. Modern ASL retains each of these meanings, as shown in examples (7), (8), (9) below. In (7) below the signer is indicating that he was blackmailing his dorm counselor after catching the counselor kissing another staff member. Now the student had the upper hand. In this sentence, CAN'T can be interpreted as indicating 'not allowed' or 'stuck'. In (8) CAN'T clearly indicates mental inability or lack of skill similar to the meaning seen in (4) above. The signer simply does not know how to read Spanish. In (9) below, the signer is discussing the rising cost of houses in his area and indicates that had he and his wife waited to purchase their home it would not have been possible. They would not have been able to. The meaning of CAN'T in (9), then, is that of a negative root possibility.

(7) role shift-up (kid) eyes up REMEMBER, TELL, role shift-down
(adult) QUIET QUIET
role shift-center (narrator) ME LUCKY [UP-TILL-NOW (from that
time on)]-top ME CONTINUE SNEAK, STEAL, STEAL FOOD
PRO.3 CANT TALK GULP
"I said "remember, don't tell," and he couldn't. I was so lucky. I
could keep on sneaking in and stealing food and he couldn't say
anything."

(8) ME LOOK-AT-lft HEART-TOUCH LOOK-AT-lft HAVE
CL-B (newspapers) ABOUT FOUR NEWSPAPERS LEFT,
MAYBE THREE.
LOOK-AT-lft ME CAN'T READ SPANISH NEWSPAPER BUT
ANYWAY (roll down window) MONEY (pull from pocket) 5
DOLLAR AMERICAN MONEY
"I was so touched by him. He had about three or four papers left, and
while I can't read Spanish I rolled down my window, and gave him
five American dollars."

(9) [NOW IF ME SHRUG-OFF(lh)]-cond
[POSTPONE (rh)]-cond MEAN ME CAN'T
HAVE HOUSE COST HIGH CRAZY
"But suppose we had waited, we wouldn't have been able to have this
house. It would have been too expensive."

It must be pointed out that while CAN'T is seen in a variety of syntactic positions in modern ASL, syntax *does* play an important role in the expression of modal notions. Shaffer (2000) suggests that, in fact, the syntactic position of the modal functions, in part, to signal either a root or epistemic reading. Further, it is suggested that if an epistemic reading is desired, the modal will only be found in clause or sentence final position, which, as I suggest in Shaffer (2000), is logical, given the topic-comment nature of the language. If, however, a root reading of the modal is desired, it will likely be found in a preverbal position, though clause final is also possible. In general, then, while root modals are found in either preverbal position or clause final position, epistemic modals are seen *only* in clause-final position.

3. A PROPOSED GRAMMATICIZATION PATH.

Following the grammaticization path described for CAN above I suggest that the physical and mental ability meanings of CAN'T are older than the negative root possibility meaning. Further I claim that there is evidence (see (10) below) of a much older meaning of CAN'T, one which indicated not *ability+ negative*, but instead one which indicated a *necessity+ negative* meaning. I suggest that this is the source for all subsequent semantic developments of CAN'T. This meaning is still retained in modern ASL and is the primary meaning for an identically articulated sign in modern LSF glossed INTERDIT and I would argue is the original meaning for both CAN'T and INTERDIT. INTERDIT in modern LSF retains its older meaning—'prohibited' or 'forbidden' and has *not* generalized to the degree to which modern ASL CAN'T has.

> (10) WE CAN'T CAN'T PLACE
> WE CAN'T CAN'T BURY HONOR
> WE CAN'T CAN'T MAKE CLEAN HERE LAND
> "We can not dedicate–we can not consecrate– we can not hallow–
> this ground."

In this sentence, (taken from a retelling of the Gettysburg Address) the sense is one of 'should not,' a meaning indicating that it would not be right to do so. The meaning of CAN'T is not one indicating the inability to do something, nor is it indicating that something is not possible. This suggests a *necessity+negative* meaning.

The *necessity+negative* meaning of CAN'T in ASL is generally interpreted in modern ASL as a denial of permission, but also is used to indicate a forbidding. While substantial evidence is lacking I hypothesize that CAN'T as we know it today is a grammaticization of MUST+ *negative* (or more precisely old LSF IL FAUT+*negative*). There is support for this claim. If one notes the older articulations of MUST (described below by Long [1918] and seen below in Figure 3) or the old LSF and modern LSF IL FAUT 'it is necessary',

one sees a striking resemblance between old IL FAUT and old CAN'T as it is articulated by McGregor and Vedetiz and others seen above.[2]

> MUST, NEED, HAVE TO—Crook the forefinger of the right "G" hand, pointing it downward and press the hand down some distance with more or less force. Sometimes the motion is repeated several times. (1918:26)

FIGURE 3. OLD LSF IL FAUT ('*it is necessary*'), MODERN LSF IL FAUT, MODERN ASL INTERDIT ('*prohibited*')/ MODERN ASL CAN'T (GIROD 1997)

I claim then, that the modern LSF sign, IL FAUT, which is translated to mean 'it is necessary' was the original basis for what in old LSF and modern LSF means 'it is necessary that...not...' and has the general meaning 'can't' in modern ASL. Utilizing the force dynamics framework as proposed by Talmy (1988) and later Sweetser (1990) we find a useful metaphor for understanding such a grammaticization path. Talmy and Sweetser suggest that we understand the expression of modal notions in terms of positive and negative forces. Applying the model to a modal such as *must*, Sweetser describes "a compelling force directing the subject towards an act" (1990:52). She notes that this interpretation of *must* differs from Talmy's. Talmy, she states, views *must* as "a barrier restricting one's domain of action to a single act" (1990:52). Sweetser would suggest that *must* has a force of order compelling one to act, rather than simply limiting or restricting one's actions (1990:52). However, both suggest understanding the expression of modal notions in metaphorical terms. Both would claim, in fact, that a system of metaphor underlies our conception of modal notions. Metaphor explains our tendency to use vocabulary "from the external (sociophysical) domain in speaking of the internal (emotional and psychological) domain" (1990:49). I suggest that this system of metaphor is at play in the understanding of MUST-*negative* concepts and their expression in ASL.[3]

---

[2] For a complete analysis of the grammaticization of MUST see Shaffer 2000.

[3] This analysis was aided by personal communications with Sherman Wilcox, Spring, 1999.

If we accept that *must* is "a barrier restricting one's domain of action to a single act" as Talmy describes it, then the sign I suggest is best described as *necessity+negative* indicates a barrier forcing one *not* to act; or, put another way, as a barrier preventing action. Analyzing the component parts, the dominant hand can be seen to be following the articulatory path for MUST or IL FAUT while the nondominant hand sets up a barrier disallowing the completion of the sign MUST and, metaphorically speaking then, disallowing the action described in the clause. Further study may reveal that the "barrier" hand may have, at an earlier time, been of a different handshape which, through a kind of symmetry change noted by Frishberg (1976), resulted in both the dominant and nondominant hand being in the index finger configuration.

Further evidence for the claim that modern ASL CAN'T grammaticized from old LSF IL FAUT comes from grammaticization theory itself. Bybee notes that it is rare to find a language with an obligation or necessity form but *without* a negative obligation or necessity form.[4] This suggests that, in fact, 1855 ASL had such a form, which during the grammaticization process underwent semantic generalization and came to be used to indicate all negative possibilities, denial of permission, *and necessity+negative*. Further, the sign has lost its transparency as MODAL+*negative* leaving what Van der Auwera (1999) calls a transparent univerbation.

4. LOSS OF TRANSPARENCY IN OTHER NEGATIVE MODALS.

The situation shown above for CAN and CAN'T is seen with the negation of virtually every other modal notion in ASL. For example, while ASL utilizes CAN'T to express the concept of *ability+negative*, a different modal form is commonly used to indicate the concept 'not possible' (glossed IMPOSSIBLE) and, in certain cases, other concepts which *could* be expressed by CAN'T. Take for example the response to the request: *pick up the piano.* Either form CAN'T or IMPOSSIBLE could be usedto respond. What is striking here is that IMPOSSIBLE appears to have grammaticizied from yet another source: not from CAN or MUST. This paradigm can be illustrated for the negation of virtually every modal notion. In each case the form differs, the resulting grammatical morpheme and its negation can not be separated, and in most cases the negation can not be identified.[5]

---

[4] Personal communication, Fall 1998.

[5] NOT-NEED appears to be the only commonly used negative modal where the negation can be identified.

5. CONCLUSION.

The claim of this paper has been two-fold. First I claim that the negative modal form CAN'T grammaticized from an older LSF form meaning 'it is necessary'. This form underwent a semantic and formational change, whereby a marker of negation was added rendering the form IL FAUT+*negative*, with the resulting meaning 'it is prohibited.' This form, while undergoing substantial semantic generaliaztion in ASL maintained its orignal conservative meaning in modern LSF. Thus, in modern LSF, the preferred reading of the sign glossed INTERDIT is 'prohibited'. Conversely, in modern ASL the sign glossed here as CAN'T is used in many semantic contexts ranging from 'prohibited' to a related meaning: denial of permission, and further to markers of inability, unavailability, and negative root possibility, and at times even, negative epistemic possibility.

A second claim of this paper is that the transparency of the negation seen in CAN'T, as well as the different source for grammaticization, will be seen for the negation of other modal forms as well. In other words, the negation of other modal concepts in ASL will yield other suppletive forms, unrelated to the modal for which they provide negation, as well as differing grammaticization paths.

REFERENCES

BROULAND, J. 1855. *Langage Mimique: Spécimen d'un Dictionaire des Signes*. Gallaudet Archives.

BYBEE, J., PERKINS R., and PAGLIUCA, W. 1994. *The Evolution of Grammar: Tense Aspect and Modality in the Languages of the World.* Chicago: University of Chicago Press.

FRISHBERG, N. 1976. *Some Aspects of the Historical Change in American Sign Language*. Ph.D. Dissertation, University of California, San Diego.

GIROD, M. 1997. *La Langue des Signes*, Tome 2. Paris, France: Edition Marketing.

LONG, JOHN. 1918. *The Sign Language: A Manual of Signs*. Washington D.C.: Gallaudet College Press.

PÉLISSIER, PAR.1856. *Iconographie des Signes*, *Tome 2*. Paris: Edition.

SHAFFER, B. 1998. The Acquisition of Markers of Modality in American Sign Language. Ms. University of New Mexico.

SHAFFER, B. 2000. *A Syntactic, Pragmatic Analysis of the Expression of Necessity and Possibility in American Sign Language*. Ph.D. Dissertation, University of New Mexico.

SMI. 1997. *The Preservation of American Sign Language*. Burtonsville, Maryland: Sign Media Inc.

SWEETSER, E. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: Cambridge University Press.

TALMY, L. 1988. Force Dynamics in Language and Cognition. *Cognitive Science,* 12.49-100.

VAN DER AUWERA, 1999. On the Typology of Negative Modals. In *Proceedings Groningen Conference on Negation*, Hoeksema E.A., Ed., pp. 25-50.

WILCOX, S. and WILCOX, P. 1995. The Gestural Expression of Modality in ASL. In *Modality in Grammar and Discourse*, J. Bybee and S. Fleishman, Eds., 135-62. Amsterdam: John Benjamins.

ENGLISH TO AMERICAN SIGN LANGUAGE MACHINE TRANSLATION OF
WEATHER REPORTS

ANGUS B. GRIEVE-SMITH

University of New Mexico

1. INTRODUCTION.

Machine translation is one of the oldest examples of applying computational technology to language. Until now, most attempts at machine translation using a signed language have been relatively word-for-word, producing sign that has none of the syntax of the target language. Not much effort has been put into producing fluent, idiomatic sign, in part because all of the attempts to date have also tried to deal with the task of sign synthesis (producing an animated, nonwritten version of the signs) at the same time.

It is possible to abstract away from the problem of sign synthesis by using a writing system. I have developed a prototype English to American Sign Language (ASL) machine translation application using Don Newkirk's (1986) Literal Orthography, a system that uses the Roman alphabet for writing signs. The translation program makes use of functionalist principles of lexical and grammatical description to produce fluent translations of National Weather Service forecasts.

I have chosen weather for the same reasons that make it a favorite testing domain of prototype natural-language processing applications: the National Weather Service genre consists primarily of a relatively small set of words and stock phrases. Albuquerque weather is particularly well-suited for this, because there is less variability than in many other cities.

2. TRANSLATING INTO ASL.

2.1. What is "Text" for a Signed Language? Translation, whether human or machine, prototypically involves reading a text in one language and writing it in another. But like !Xóõ, Fulani, and a large number of other spoken languages, ASL has no standard written form. Unlike these languages, ASL is so different from the prototypical written language that the possibility of writing it has not occurred to most of its speakers.

Because of this, sign synthesis and machine translation are often confused. Many of the prototype sign synthesizers (e.g. Ohki, Sagawa, Sakayama, Oohira, Ikdea and Fujisawa 1994) have treated synthesis as an integral part of machine translation. This is different from the spoken language situation: we expect a French-English translator to produce written English, not synthesized speech. We would even expect a French-Fulani translator to produce written, not spoken, Fulani, even though very few people write Fulani.

For demonstration purposes, I have adopted a short-term solution, to use Newkirk's (1986) "Literal Orthography" one of the nine writing systems that have

been developed for signed languages. The system is based on the Roman alphabet, which allows easy integration with ASCII-based Unix systems. It is more complete than ASCII-Stokoe (Mandel 1993), the other ASCII-based writing system, and can nearly represent the full range of lexical and classifier signs of ASL. The following sign can be translated as "wind" or "windy" in English:

(1)   so-bles

The prefix "so" refers to the hands making "contrary movement" meaning that while one hand (the dominant hand, in this case) is moving away from its side, the other is moving towards its side. "bl" indicates that the handshape is the one used for the number 5 in the American Manual Alphabet. "e" means that the hands move towards the non-dominant side of the body, and "s" means that the movement is reduplicated.

The system as Newkirk developed it in 1986 did not represent facial and body gestures or fingerspelling, two integral components of ASL, so I developed ad hoc conventions for representing these. The facial and body gestures are indicated by a short word following the word for a lexical or grammatical sign, that can be easily distinguished from a sign by the number and combination of letters used. Examples of this notation is given below:

(2)   a. so-bles w
        b. so-bles r

Example (2a) indicates that the word "so-bles" is modified by the diminutive facial adverb "w" (pursing the lips) to translate the English "breezy." The English word "gusty" is translated with "so-bles" modified by the intensifying adverb "r" (squint) in (2b).

Fingerspelling is simply represented by an "@" sign placed before the Roman version of the fingerspelled word, as in the following example for "miles per hour":

(3)   @mph

This is a workable solution for the short term. In the long term there are two possible solutions. One is the adoption of a writing system by the American Deaf Community, which would have other desirable effects such as greater status for ASL and increased literacy. Another possibility is integration with an eventual sign synthesis application.

2.2. Corpus Planning. With translation, as with any instance of language contact, issues of power come into play. What happens when a translator works with a topic that is not usually discussed in the target language? Two possibilities exist: that the translator could borrow ad hoc from the source language or another language of prestige, or that someone could invent new words to be used for this topic. The first possibility involves a loss of prestige for the language, since it is deemed incapable of

expressing this topic. The second is a more grassroots procedure, and allows for community involvement in the development of the language.

A well-known example of this is that of computer terminology in French. Since most computer equipment was invented either in the United States or in Japan, there were no French words for concepts like "software" and "the Web." In the case of "the Web," the word "web" has been borrowed from English, while in the case of "software," a group of French community leaders decided to invent a word, "logiciel," which has been adopted by French speakers and even clipped to "log" on occasion. A coinage may be rejected by the users of the language; for example, the invented word "hambourgeois" has been ignored in favor of the English borrowing "hamburger."

The situation for weather description in ASL is similar. Deaf people do talk about the weather, so the conversational vocabulary exists already. What is not available, according to informal discussions with native signers and interpreters, is a formulaic register of jargon and expressions comparable to that used by the National Weather Service, as seen in example (4) from January 20, 1999:

(4)   Tonight: Mostly cloudy with a slight chance of rain showers.
        Breezy and mild. West to southwest winds 10-20 mph.

The short-term solution adopted for this project has been informal consultation with experts who are native speakers of ASL, to determine the conversational vocabulary used to discuss weather and adapt it for the translation of National Weather Service forecasts. A long-term solution would involve corpus design by a group of interested community leaders.

3.   THE APPLICATION.

As is common with machine translation systems, the application consists of four components: a lexical analyzer, a parser, a transfer module and a generation module. In addition, there is an initial module that obtains the weather reports from the World Wide Web. Several of the components use freely available Perl modules, packages designed to assist in those particular tasks for spoken or computer languages.

3.1. Retrieval and Lexical Analysis of Weather Reports in English. The weather reports are published in English by the National Weather Service on the World Wide Web at <http://iwin.nws.noaa.gov>. Geo::WeatherNOAA is a Perl module, developed by Mark Solomon to assist in downloading the reports from the National Weather Service site and available from the Comprehensive Perl Archive Network (CPAN) at <http://www.cpan.org>. The report for January 20, 1999, as produced by Geo::WeatherNOAA looks like this:

(5) Today: Partly cloudy. Increasing west to southwest winds 15-25
mph this afternoon.
Tonight: Mostly cloudy with a slight chance of rain showers.
Breezy and mild. West to southwest winds 10-20 mph.
Thursday: Windy and slight cooler with a chance of rain
showers. West to northwest winds increasing to 20-30 mph with
higher gusts.

Parse::Lex is a lexical analysis module developed by Philippe Verdret and available from CPAN. It allows the user to define an arbitrary set of lexical categories (for example, part-of-speech categories), and then apply those categories to a given text. For this application, I chose to use ad hoc semantically-based categories instead of syntactic categories such as part of speech. Parse::Lex was configured to tag the text with the lexical category for each word; for example, "today" was tagged with "<day>."

Pawley and Syder (1983) argue that nativelike fluency in a language can best be explained by postulating that the lexicon is composed of chunks that can be larger than typical words or morphemes. I decided to configure the lexical analyzer based on these principles, so any sequence of words that seemed to be a set phrase in National Weather Service terminology was tagged as a single word; for example, "rain showers" was tagged as a single unit with "<precip>." For this pilot study, the chunks were chosen based on my intuitive judgments as a native English speaker; in the future, this lexical analysis could be based on actual counts of token frequency within National Weather Service texts.

Example (6) shows our forecast from January 20, 1999, after being tagged by the lexical analyzer:

(6) today <Day> : <Punc> partly cloudy <Sky> . <Punc> increasing <Change> west <Direct_Locat> to <Preposition> southwest <Direct_Locat> winds <Wind> 15-25 <NumRange> mph <Degree> this <Demonstrative> afternoon <Time> . <Punc> tonight <Time> : <Punc> mostly cloudy <Sky> with <Preposition> a <Determiner> slight <Degree> chance <Chance> of <Preposition> rain showers <Precip> . <Punc> breezy <Sky> and <Conjunction> mild <Heat> . <Punc> west <Direct_Locat> to <Preposition> southwest <Direct_Locat> winds <Wind> 10-20 <NumRange> mph <Degree> . <Punc> thursday <Day> : <Punc> windy <Sky> and <Conjunction> slight <Degree> cooler <Heat> with <Preposition> a <Determiner> chance <Chance> of <Preposition> rain showers <Precip> . <Punc> west <Direct_Locat> to <Preposition> northwest <Direct_Locat> winds <Wind> increasing <Change> to <Preposition> 20-30 <NumRange> mph <Degree> with <Preposition> higher <Degree> gusts <Wind> . <Punc>

   3.2. Parsing English Weather Reports. The tagged reports are then parsed using Parse::RecDescent, a Perl module developed by Damian Conway of Monash University, and available from CPAN. RecDescent is a recursive descent parser that can be configured to produce a parse tree based on an input text in a particular language. Since the weather reports are formulaic and do not rely on complicated syntactic structures, it was possible to directly create a semantic representation, without the intermediate step of a syntactic tree.

   Every National Weather Service forecast can be divided into four semantic components, corresponding to the real-world weather domains of sky, precipitation, wind and heat. This division is often reflected in the syntactic structure of the reports: often each of the components has its own sentence. Within each component, the structures can be as simple as a single lexical item ("partly cloudy"), or more complicated, with phrases specifying probability, wind speed or change of state. These were all reflected in the parser. What follows is an excerpt from the parse tree for the January 20, 1999 forecast, showing a "wind phrase" for the first sub-forecast:

(7)   [1][1]{windP}{time}[1]{day}=
        [1][1]{windP}{time}[1]{time}=afternoon
        [1][1]{windP}{time}[1]{prep}=
        [1][1]{windP}{time}[1]{adv}=
        [1][1]{windP}{degree}=
        [1][1]{windP}{conj}=
        [1][1]{windP}{direct1}=west
        [1][1]{windP}{direct2}=southwest
        [1][1]{windP}{main}=winds
        [1][1]{windP}{change}=increasing
        [1][1]{windP}{speedP}{total}=15 - 25
        [1][1]{windP}{speedP}{lo}=15
        [1][1]{windP}{speedP}{hi}=25
        [1][1]{windP}{locatP}[1]{main}=
        [1][1]{windP}{locatP}[1]{degree}=
        [1][1]{windP}{locatP}[1]{prep}=
        [1][1]{windP}{locatP}[2]{main}=
        [1][1]{windP}{locatP}[2]{degree}=
        [1][1]{windP}{locatP}[2]{prep}=
        [1][1]{windPlus}{total}=
        [1][1]{windPlus}{degree}=
        [1][1]{windPlus}{main}=
        [1][1]{windPlus}{prep}=
        [1][1]{windPlus}{speed}=
        [1][1]{windPlus}{locatP}[1]{main}=
        [1][1]{windPlus}{locatP}[1]{degree}=
        [1][1]{windPlus}{locatP}[1]{prep}=

The complexity of this tree requires a short explanation. Each of the lines represents a piece of weather information that might be found in the phrase of a forecast describing the wind for a particular time (in this case time [1][1], or the first time mentioned in the first paragraph of the forecast). The lines with nothing after the = sign represent information that was not found in the input forecast. The information that was found were time (afternoon), direction (west-southwest), change (increasing), and speed (15-20 mph).

3.3. Transfer of the Parse Tree into ASL. The transfer into an ASL-based tree is accomplished by a simple lookup table. The table is based on the items in the lexical analyzer, each with a translation in idiomatic ASL, written in Newkirk notation. I developed a Perl script to read each line of the English parse tree and look up the word in the table. It then recreates the parse tree, replacing the English with the ASL. The following example shows the transferred version of the excerpt above:

(8) [1][1]{windP}{time}[1]{day}=
[1][1]{windP}{time}[1]{time}=byy:yd
[1][1]{windP}{time}[1]{prep}=
[1][1]{windP}{time}[1]{adv}=
[1][1]{windP}{degree}=
[1][1]{windP}{conj}=
[1][1]{windP}{direct1}=woo:o
[1][1]{windP}{direct2}=sooa woo:o
[1][1]{windP}{main}=so-bles
[1][1]{windP}{change}=so-bray c
[1][1]{windP}{speedP}{total}=15 - 25
[1][1]{windP}{speedP}{lo}=15
[1][1]{windP}{speedP}{hi}=25
[1][1]{windP}{locatP}[1]{main}=
[1][1]{windP}{locatP}[1]{degree}=
[1][1]{windP}{locatP}[1]{prep}=
[1][1]{windP}{locatP}[2]{main}=
[1][1]{windP}{locatP}[2]{degree}=
[1][1]{windP}{locatP}[2]{prep}=
[1][1]{windPlus}{total}=
[1][1]{windPlus}{degree}=
[1][1]{windPlus}{main}=
[1][1]{windPlus}{prep}=
[1][1]{windPlus}{speed}=
[1][1]{windPlus}{locatP}[1]{main}=
[1][1]{windPlus}{locatP}[1]{degree}=
[1][1]{windPlus}{locatP}[1]{prep}=

Here we see the same parse tree, but with all the individual words translated into ASL. For example, the English "afternoon" becomes the ASL "byy:yd," "southwest" becomes "sooa-woo:o," and so on.

3.4. Generation of Fluent ASL. The ASL generation module uses the notion of "sentence stems" proposed by Pawley and Syder (1983:205-215) to generate fluent written ASL. The Perl script first takes an inventory of the kinds of information present in the semantic representation, and generates a formulaic phrase for each one. These formulas all use ASL grammar, including topic-comment structure and nonmanual grammatical morphemes. The content that is output by the transfer module is then plugged into the formulas, producing fluent written ASL. The translated weather report for January 20, 1999 thus looks like this:

(9) s-yya b so-blhihyeuws r. woo:o sooa woo:o so-bles so-bray c 15
25 @mph byy:yd.
byyhayri b so-blhihyeuws rm. si-byays si-seeis s-blhasas w. so-
bles w. aeexion m. woo:o sooa woo:o so-bles 10 20 @mph.
nyms b so-bles. s-blix's areeixx' w. si-seeis s-blhasas w. woo:o
hooy woo:o so-bles so-bray c 20 30 @mph.

4. FUTURE WORK.

There are several possibilities for the extension of this work. The output needs to be cross-checked with a native signer to ensure that it is indeed fluent, idiomatic ASL. A double-blind method could be employed, whereby a signer who knows the Literal Orthography reads the translation of a randomly chosen forecast, and a native-speaker evaluator is then tested on their understanding of the translation.

The lexical analyzer and parser are still not completely adjusted to the full range of weather reports in English; since much of the training was done during a mild winter in Albuquerque, words for snow have only recently been added to the lexicon. A frequency analysis of the corpus would determine whether the "chunks" currently used in the lexicon correspond to chunks used in these reports. Additional corpus planning should eventually be undertaken with community leaders to develop standard ASL phrases corresponding to the set expressions of the National Weather Service.

5. CONCLUSION.

The production of ASL by this prototype system shows that machine translation from English into ASL is feasible. The process is relatively straightforward, if we focus on the translation itself and leave aside synthesis issues. There are several projects underway to produce both user-friendly writing systems for signed languages, and sign synthesis applications, so machine translation that is usable by most signers seems likely to appear soon.

REFERENCES

MANDEL, MARK A. 1993. ASCII-Stokoe notation: A computer-writeable transliteration system for Stokoe Notation of American Sign Language. Manuscript. Available at <http://world.std.com/~mam/ASCII-Stokoe.html>.

NEWKIRK, DON E. 1986. Outline of a proposed orthography of American Sign Language. Manuscript. Available at <http://users.home.net/dnewkirk/signfont/orthog.htm>.

OHKI, MASARU, HIROHIKO SAGAWA, TOMOKO SAKIYAMA, EIJI OOHIRA, HISAHI IKEDA AND HIROMUCHI FUJISAWA. 1994. Pattern Recognition and Synthesis for Sign Language Translation System. *Assets* 10.1-8.

PAWLEY, ANDREW, AND FRANCES HODGETTS SYDER. 1983. Two Puzzles for Linguistic Theory: Nativelike Selection and Nativelike Fluency. *Language and Communication*, ed. by Jack C. Richards and Richard W. Schmidt, 192-225. New York: Longman.

ON MANAGING COMPLEXITY IN SIGN LANGUAGE RECOGNITION[*]

CHRISTIAN VOGLER AND DIMITRIS METAXAS

Vision, Analysis, and Simulation Technologies Laboratory

University of Pennsylvania

1. Introduction.

A pivotal problem in American Sign Language (ASL) recognition is managing complexity. Because of its incredibly rich set of inflections, that causes signs to appear in many different forms, it is futile to model the language on a per-sign basis for large-scale recognition. Instead, it is necessary to break down the signs into their constituent phonemes, which are limited in number.

Such a breakdown, however, raises two interesting problems: First, current work in ASL phonology still has some gaps that make it difficult to apply directly to ASL recognition. Second, many events occur simultaneously in ASL, such as the movements of the strong and the weak hands, or a change of handshape, and a movement of the arm.

To solve the first problem, we experimented with a modified version of Liddell and Johnsons's Movement Hold Model. We show what changes were necessary to adapt this model to computer ASL recognition. The second problem is serious, because existing recognition frameworks are sequential in nature and thus inadequate to capture phonemes that appear simultaneously in a sign. We discuss a modification to our recognition framework that models the strong and the weak hands as moving in parallel, independently from each other. The advantage of this approach is that it does not need to consider all possible combinations of phonemes *a priori.*

Imagine a future in which linguists would never have to prepare a time-consuming transcription of sign language data again. Instead, they could simply instruct a computer to prepare the transcription from video. Admittedly, this future is still a way off. But this potential eventuality captures the essence of sign language recognition very well. In its simplest form, sign language recognition consists of converting information from the movement of the arms, the hands, the face, and the rest of the body of the signer into a human- or computer-readable transcription. There are several other problems related to sign language recognition, among them the estimation of the three-dimensional movement parameters of the signers body parts from two-dimensional video images, and the translation of the sign language transcription to English or another language. In this

paper we are concerned only with the actual recognition process, that is, converting the movement parameters into a transcription.

Besides making tedious transcriptions easier, sign language recognition has several other applications: applying different approaches to modeling sign language with a computer for recognition purposes could help toward clarifying sign language linguists' positions. For example, there is considerable disagreement among linguists about how to describe sign language phonology. Testing which phonological models work well for sign language recognition could contribute toward resolving these disagreements. In addition, together with a machine translation system, sign language recognition could make deaf-hearing interaction easier, especially in the absence of sign language interpreters. There is also a strong link between sign language recognition and gesture recognition. The latter is an important research topic for future human-computer interaction because speech and gestures constitute the most natural form of expression for most humans. Because sign language is a full-fledged language, there are constraints on the appearance of the language and, thus, many problems may be easier to solve in sign language recognition, before they are applied to more general gesture recognition.

In this paper, we specifically discuss the relationship between sign language linguistics and sign language recognition. Most previous work in this field has more or less ignored linguistics; however, as the field advances and the vocabulary of recognizers becomes larger, new work cannot afford to ignore linguistics anymore. The main reason for this is computational complexity. Consider as a case in point that American Sign Language (ASL) has a mapped vocabulary of approximately 6,000 signs, of which many are highly inflected. With all possible inflections factored in, the number of distinct forms that can appear in ASL is several orders of magnitude higher than 6,000. If we ignore linguistics, we would have to model all these distinct forms explicitly with a computer, which is impossible. Computers are not powerful enough to perform the computations that are necessary for recognition over so many different forms. There are not enough resources available to program so many forms into a computer either.

Sign language phonology can help to reduce the computational complexity of sign language recognition, because there are many fewer different phonemes than there are signs. If signs are broken down into their constituent phonemes, we have to worry only about capturing the possible combinations of phonemes, instead of all possible combinations and appearances of signs. However, even if a recognizer uses ASL phonology, there are still two problems to be resolved. First, there are several conflicting models of ASL phonology (Stokoe 1960, Liddell and Johnson 1989, Brentari 1995, Sandler 1989, Coulter 1993) and it is not clear which one is most appropriate for sign language recognition. Second, ASL has both sequential and parallel contrast; that is, phonemes can appear both sequentially and in parallel. Parallel contrast poses a problem because we have to worry about all possible combinations of phonemes that can occur in parallel. With the conservative estimate that there are roughly 30 handshapes, 20 major body locations, 8 hand orientations, 40 types of movements, and two hands, there are approximately $(30 \times 20 \times 8 \times 40)^2 = 3.8 \times 10^{10}$ theoretically possible combinations. Of course, not all combinations occur in ASL, but even so, the number is so high that it is

impossible to consider all possible combinations of phonemes *a priori* in a recognizer. Instead, a recognizer must be able to handle possible combinations on the fly.

To tackle the first problem, we discuss the use of Liddell's and Johnson's (1989) Movement-Hold phonological Model for ASL recognition, and what modifications we had to make to get it to work within a Hidden Markov Model (HMM) statistical framework (Rabiner 1989). To tackle the second problem, we assume that phonemes occurring simultaneously are independent from one another. More specifically, we assume that the strong and the weak hands move independently from each other. In doing so within a statistical framework, it is possible to combine the strong and the weak hands easily by multiplying their joint probabilities. The question is whether this independence assumption is valid. We provide some experimental evidence that seems to show that this assumption is at least partially valid in the context of ASL recognition.

In the following sections, we discuss related work, then give an introduction to the Movement-Hold phonological Model. Then we discuss what modifications we had to make to this model and how to incorporate the simultaneous aspects of ASL phonology. Finally, we provide experimental evidence to back up our claims.

2. RELATED WORK.

Starner and Pentland (1995) use a view-based approach with a single camera to extract two-dimensional features as input to HMMs with a 40-sign vocabulary and a highly constrained sentence structure consisting of a pronoun, verb, noun, adjective, and pronoun in sequence. They make the assumption that whole signs are the smallest unit in ASL and otherwise do not mention any linguistic aspects.

Liang and Ouhyoung (1998) use HMMs for continuous recognition of Taiwanese Sign Language with a vocabulary between 71 and 250 signs. Their work is based on Stokoe's model (Stokoe 1960), which, however, has been superseded by other phonological models such as (Liddell and Johnson 1989). Unlike other work in this area, they integrate the handshapes, positions, orientations, and movements of signs via high-level computational processing, instead of the HMM level.

Hienz and colleagues (1999) use HMMs for continuous video-based sign language recognition of German Sign Language with a 52-sign vocabulary. In addition, they use stochastic unigram and bigram language models over the possible sequences of signs to improve recognition performance.

We previously used HMMs to model phonological aspects of ASL (Vogler and Metaxas 1997, 1999a) with an unconstrained sentence structure. The Movement-Hold phonological model (Liddell and Johnson 1989) formed the basis for our work. In (Vogler and Metaxas 1999b), we introduced a framework for modeling simultaneous aspects of ASL for the first time and we gave experimental evidence for modeling the strong and weak hands in parallel. Whereas our previous paper discussed the technical aspects of using ASL phonology for recognition, this paper discusses some of the linguistic aspects.

3. THE MOVEMENT-HOLD MODEL.

The earliest phonological description of ASL (Stokoe 1960) provided only for simultaneous contrast and dismissed sequential occurrences of phonemes as insignificant. Liddell and Johnson (1989) argued convincingly against the assumption that sequences of phonemes were noncontrastive. In fact, sequential contrast forms the basis of their Movement-Hold Model. Models that emphasize sequential contrast are interesting for sign language recognition, because the Hidden Markov Model (HMM) statistical framework is sequential by nature. It is much easier to use this framework to handle sequences of events (i.e., phonemes) than events that happen simultaneously. We now give an overview of the parts of the Movement-Hold Model that we used for our recognizer. This overview is by no means exhaustive. Also, we do not consider other sequential phonological models (Brentari 1995, Sandler 1989) at the moment, because the whole idea of phonological modeling is new in the field of sign language recognition. Obviously, in future work it will be necessary to test the merits of using other phonological models.

3.1. Characteristics of the Movement-Hold Model. There are two major classes of segments in the Movement-Hold Model (Liddell and Johnson 1989), which are called MOVEMENTS and HOLDS. Movements are defined as those segments during which some aspect of the signer's configuration changes, such as a change in handshape, a hand movement, or a change in hand orientation. Holds are defined as those segments during which all aspects of the signer's configuration remain stationary; that is, the hands remain stationary for a brief period of time.

Signs are made up of sequences of movements and holds. Some common sequences are *HMH* (a hold followed by a movement followed by another hold, such as 'good'), *MH* (a movement followed by a hold, such as 'sit'), and *MMMH* (three movements followed by a hold, such as 'father'). Attached to each segment is a BUNDLE OF ARTICULATORY FEATURES that describes the hand configuration, orientation, and location. In addition, movement segments have features that describe the type of movement (straight, round, sharply angled), as well as the plane and intensity of movement. See **FIGURE 1** for a schematic example:

FIGURE 1. SCHEMATIC EXAMPLE OF THE SIGN FOR 'FATHER' IN THE MOVEMENT-HOLD MODEL

M straight back	**M** straight forward	**M** straight back	**H**
near forehead	touches forehead	near forehead	touches forehead
5-hand	5-hand	5-hand	5-hand
points up	points up	points up	points up
faces left	faces left	faces left	faces left
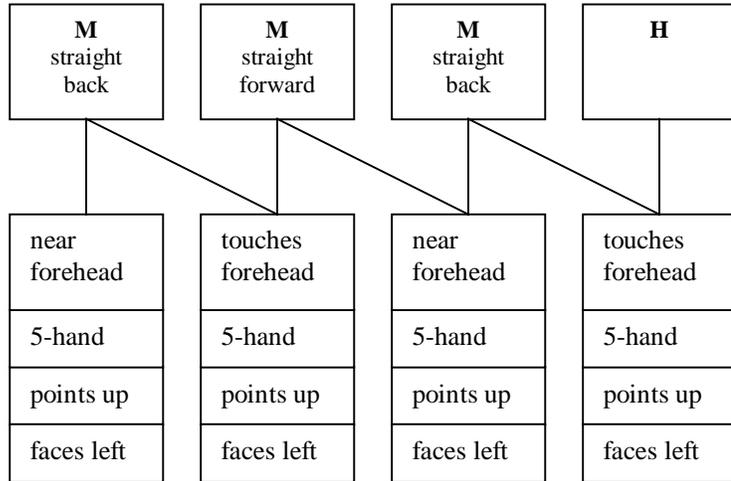
The Movement-Hold Model also makes a distinction between global movements and local movements. Global movements are straight, circular, or angular movements, and they are described directly in the movement segment, as show in FIGURE 1. Local movements are usually superimposed on the global movements and consist of wrist rotations, zigzag movements, and wavy movements, among others. These are described in the articulatory features attached to both movement and hold segments.

In **TABLE 1** and Figure 2, we give a partial overview of the different descriptions of movements and locations that we used in our work. Note that they sometimes differ from Liddell and Johnson's descriptions for reasons explained in the next section. In addition, the locations can be modified with the distance from the body and with the vertical and horizontal distance from the basic location.

TABLE 1. PARTIAL LIST OF MOVEMENTS. (Note that the description of the movements deviates from the approach used by the Movement-Hold Model.)

Type of Movement	Transcriptions used
straight	str_{Away}, str_{Toward}, str_{Down}, str_{Up}, str_{Left}, str_{Right}, $str_{DownAway}$, $str_{DownRightAway}$
short straight	$Str_{ShortUp}$, $str_{ShortDown}$
circle in vertical plane	rnd_{VP}
wrist rotation	rot_{Away}, rot_{Toward}, rot_{Up}, rot_{Down}

FIGURE 2. PARTIAL LIST OF BODY LOCATIONS USED IN THE MOVEMENT-HOLD MODEL

iFH FH %iFH

NS

MO

iCN CN %iCN

ST %iST

CH

TR

ABu

AB
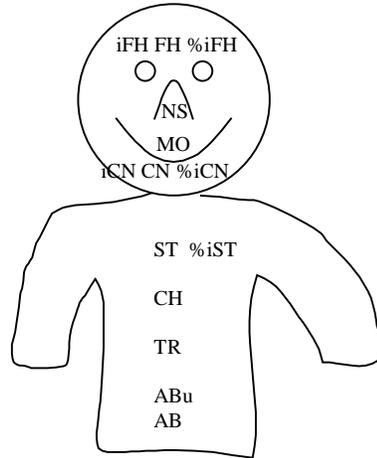
If a location does not touch the body, it can be prefixed with one of these distance markers: *p* (proximal), *m* (medial), *d* (distal), or *e* (extended), in order of increasing distance from the body. If a location is centered in front of the body, the distance marker is suffixed with a 0. If the location is at the side of the chest, the distance marker is suffixed with a 1, and if the location is to the right (or left) of the shoulder, the distance marker is suffixed with a 2. For example, *d-1-TR* means a location a comfortable arm's length away from the right side of the trunk (torso).

Further markers describe the vertical offset to the basic location and whether the location is on the same side or opposite side of the body as the hand. These are described in detail in (Liddell and Johnson 1989).

3.2. Application of the Movement-Hold Model to Recognition. We would like to use the same classification scheme as the Movement-Hold Model for segments with our recognizer. From a technical point of view, it makes sense to make an even sharper distinction between the two classes of segments than the original Movement-Hold Model does. The reason is that the estimation of locations is very different from the estimation of movements. In the former case, we are interested in the positional parameters of the hands, whereas the velocity of the hands should be zero. In the latter case, we are interested in the velocity and in the trajectory of the hands, whereas the position of the hand could be anywhere. By making this sharp distinction between the two classes of segments, we can make this difference in estimating the parameters explicit and increase recognition accuracy.

3.2.1. Local and global movements. This sharp distinction between the two classes of segments, however, leads to an immediate problem. What about local movements, which are described in the articulatory features and thus can occur in both movement and hold segments? There are some signs that exhibit *only* these local movements, such as the sign for 'interpret.' It consists of three hold segments, with a series of wrist rotations—forward, backward, and forward again—attached to them. We would like not to have any motion at all during the hold segments for the reasons given before, but we cannot ignore the wrist rotation either, because it is the most prominent feature of the sign. On the other hand, sometimes a wrist rotation is superimposed over another movement, such as in the sign for 'try,' which consists of a symmetrical downward and outward movement of both hands, while the wrists rotate counterclockwise. In this case, the downward and outward movement of the hands is by far the most prominent feature of the sign.

Because of these problems, we maintain that putting global movements in the movement segments, and putting local movements in the articulatory features is an artificial distinction. At least for sign language recognition, it makes more sense to put both local and global movements in the movement segments, instead of in the articulatory features. In addition, we distinguish between formerly local movements that constitute the only kind of movement in the course of a sign, such as in 'interpret,' and local movements that are superimposed over other movements, such as in 'try.' We reclassify the former type of local movement as global. It is for this reason that wrist rotations appear in **TABLE 1**. Hence, the transcription of the sign for 'interpret' changes from the *HHH* pattern to the *MMMH* pattern; that is, the three wrist rotation movements, followed by a final hold.

3.3.2. Short and long movements. Another problem with the application of the Movement-Hold Model to sign language recognition arises for purely technical reasons. In the Movement-Hold Model the manner of a movement can be modified with such attributes as long, short, tense, and so on. For example, a straight movement can be short, long, or normal. Ideally, we should represent the different lengths of straight movements with the same underlying model of what a straight movement should look like.

In practice, however, it is difficult to use the same underlying representation, because the physical appearance of a short straight movement is very different from the appearance of a normal straight movement. Both the velocities and the trajectories look very different. For this reason, we tentatively model short movements and normal movements as distinct classes, as shown in **TABLE 1**. However, making this distinction is not a good long-term solution, because it increases the number of phonemes needlessly and thus increases computational complexity. Future research should look into using the same underlying representation for all movement modifications.

3.2.3. Problems with the MMMH pattern. A third problem with the application of the Movement-Hold Model to sign language recognition arises for both technical and linguistic reasons. Recall that we would like to make a sharp distinction between movements and holds, which implies that movements should be location-independent. In other words, we can estimate the locations of signs only during hold segments. The *HMH*

pattern causes no problem, because the location can be estimated at both the beginning and the end of the sign. The *MMMH* and to a lesser extent the *MH* pattern, however, cause problems, because they start with one or more movement segments, and the hold segment appears only at the end of the sign. Hence, we would be able to estimate the locations only at the end of the signs, even though locations are also significant at the beginning of the signs.

Therefore, it is necessary to estimate the locations at the beginning of the signs, even if they do not start with a hold segment. This problem looks like a technical one on first glance, but a closer look reveals that it also stems from an inconsistency in the Movement-Hold Model. Movement segments are by definition those segments in which aspects of the configuration are *changing*. Yet, the bundles of articulatory features describe *static* aspects of the signs. Why, then, are these bundles attached to movement segments, as shown in **FIGURE 1**?

To solve this inconsistency, we introduce a new type of segment that does not have an equivalent in (Liddell and Johnson 1989).[1] We call this new segment "X". It is very similar in function to a hold segment in the sense that it also captures static aspects of a sign. The hand has to remain stationary for a brief amount of time in an hold segment, but we do not make this requirement for the X segment. Its only function is to capture a 'snapshot' of the static parameters of the sign, even if the hands are currently in motion. We use the X segment in the beginning of a sign, whenever it previously started with a movement segment. That is, *MMMH* becomes *XMMMH*, and *MH* becomes *XMH*. We remove all attachments of the feature bundles to the movement segments and attach them to the X segments instead, as shown in Figure 3:

---

[1] In the latest, as of yet unpublished version of the Movement-Hold Model, there is also an X segment, which seems to be very similar to the X segment that we use in this paper.

FIGURE 3. USE OF THE X SEGMENT IN THE SIGN FOR 'SIT' (STRONG HAND)

X	M straight down	H
neutral space		touches left hand
H-hand		H-hand
forward		forward
down		down

.
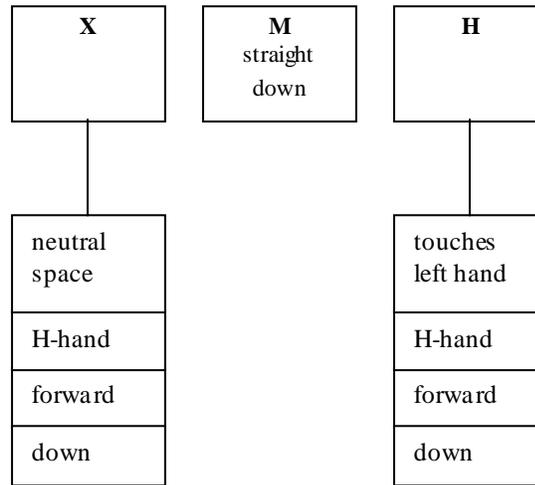
3.3. Movement Epenthesis. Liddell and Johnson 1989 also describe several phonological processes in ASL. In order to achieve robustness, a recognition system must be able to cope with such processes. The most basic, and at the same time also most important phonological process in ASL, is called *movement epenthesis*. It consists of the insertion of extra movements between two adjacent signs and it is caused by the physical characteristics of sign languages. For example, in the sequence 'father read,' the sign for 'father' is performed at the forehead and the sign for 'read' is performed in front of the trunk. Thus, an extra movement from the forehead to the trunk is inserted that does not exist in either of the two signs' lexical forms (Figure 4):

FIGURE 4. MOVEMENT EPENTHESIS. (The arrow in the middle frame indicates an extra movement between the signs for 'father' and 'read' that is not present in their lexical forms.)

Movement epenthesis poses a problem for ASL recognizers, because the appearance of the movement depends on which two signs appear in sequence. We handle this problem by modeling such movements explicitly. We do not yet model other phonological processes, such as metathesis, but future work should handle them, as well.

3.4. Parallel Aspects of ASL. A look at the feature bundles in **FIGURE 1** and Figure 3 reveals that many of the phonemes occur in parallel, such as the handshape, and the location. Also, many signs are two-handed, so it is necessary to consider both hands moving in parallel. Unfortunately, it is not obvious how a sign language recognizer will be able to handle the parallel aspects.

3.4.1. The problem: complexity. The heart of the problem is that the Hidden Markov Model (HMM) statistical framework is a sequential framework by nature. In essence, it computes the probabilities that a sequence of signs or phonological segments matches the motion parameters of the signer's body; see (Vogler and Metaxas 1997, 1999a) for details. For this reason, all phonemes that occur in parallel within a segment must be combined into a single whole, so as to be considered during the recognition process. If such a combination into a whole were to take place, it would require us to model all possible combinations *a priori,* before ever using the recognizer.

Modeling all possible combinations *a priori* is impossible, because there are far too many combinations of phonemes that can occur in sign language. If we assume that there are roughly 30 handshapes, 20 major body locations, 8 hand orientations, 40 types of movements, and two hands in ASL, there are approximately $(30 \times 20 \times 8 \times 40)^2 = 3.8 \times 10^{10}$ theoretically possible combinations. An order of magnitude of 10^{10} exceeds both the storage capacity and the speed of present-day computers and will do so for the foreseeable future. In addition, each phoneme combination that is modeled *a priori* would require several real-world examples of data, so that the HMM framework can be trained

on these data. Obtaining such data examples for all possible combinations surely exceeds the resources of the most ambitious research project.

Thus, it becomes clear that considering all possible combinations of phonemes makes the recognition task too complex. We now discuss a method of reducing the complexity of modeling the parallel aspects of ASL.

3.4.2. Proposed solution: independent channels. Because *a priori* modeling of the parallel aspects of ASL makes the recognition task too complex, we need to use a method that avoids the such modeling in favor of putting together the phoneme combinations on the fly. The statistical nature of the HMM framework suggests such a solution: assume that the different phonemes can occur independently from one another. That is, assume that the handshape within a sign is independent from the hand movement within a sign; or assume that the strong hand and the weak hand move independently from each other. More specifically, we assume that the two hands move in independent channels, as well as the handshape and orientation parameters. See FIGURE 5 for a schematic example of the sign for 'sit':

FIGURE 5. INDEPENDENT CHANNELS FOR THE STRONG AND WEAK HANDS IN THE SIGN FOR 'SIT'.
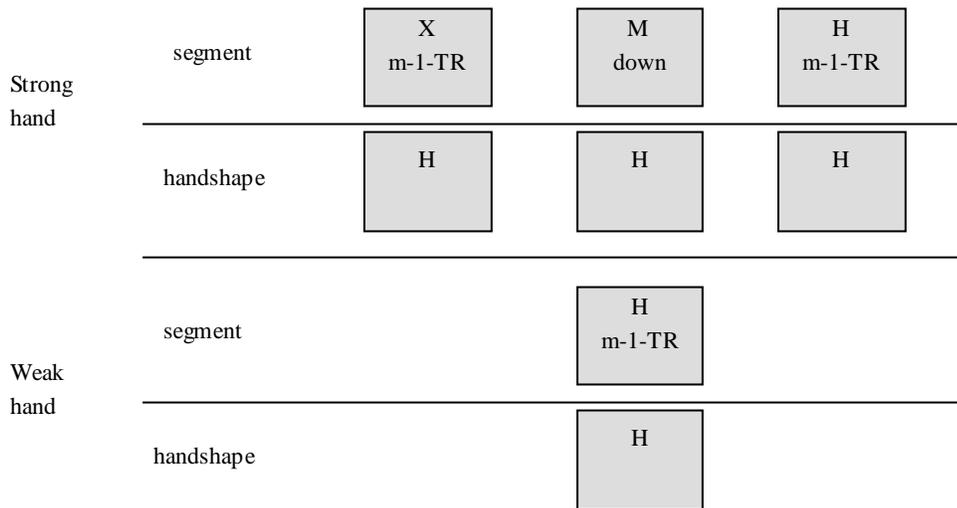
The independence assumption allows us to combine parallel phonemes on the fly simply by multiplying the probabilities that each individual phoneme occurred in a given sequence of signs.

From a computer science point of view, this assumption is a very elegant way to reduce computational complexity. From a linguistic point of view, however, it leaves something to be desired. Although Liddell and Johnson (1989) present an argument that

the two hands should move relatively independently from each other, it is not clear that the independence assumption is valid. Do the two hands really move independently from each other? Likewise, do the handshapes really occur independently from the hand movements?

On the one hand, the presence of two-handed signs where the hands move symmetrically or in an alternating pattern suggests that the answer to these questions is no. Conversely, when the weak hand anticipates the starting position of a two-handed sign, while the strong hand still signs a one-handed sign, it seems to have some freedom in when and how fast it moves to the starting position. Hence, the two hands may at least be partially independent from each other. The bottom line will be how well the independence assumption works in sign language recognition, and not necessarily how well it fits with sign language recognition. Our experiments in the next section suggest that this independence assumption does help to improve recognition rates.

4. EXPERIMENTS.

We ran three types of experiments that tested the relative merits of using whole signs, breaking down whole signs into phonemes according to our modified version of the Movement-Hold Model using the strong hand only, and modeling both the strong hand and the weak hand independently. Our sign language database consisted of 499 sentences with a 22 sign vocabulary. The sentences were constrained only by the vocabulary and the grammar of ASL. We collected our sign language database with an Ascension Technologies MotionStar^{TM} magnetic tracking systems, which gave us the 3D positions and velocities of the arms and the hands.

Of these 499 sentences, we used 400 sentences to train the HMM framework and used the other 99 sentences to test recognition accuracy. To keep the experiments simple, we looked only at the positions and velocities of the hands, not at the handshapes, orientations, and nonmanual markers. The technical details of the experiments can be found in (Vogler and Metaxas 1999a, 1999b).

The results of the experiments are given in FIGURE 6. We use word accuracy as our evaluation criterion, which is computed by subtracting the percentage of erroneously inserted signs from the number of correctly spotted signs in the test sentences.

The results are revealing in two ways. First, there is no significant difference in recognition accuracy between using whole signs and using phonemes as the basic unit in sign language. This result is encouraging, because phoneme modeling has much lower computational complexity than whole-sign modeling, and hence it will scale better to larger vocabularies. Second, the independent modeling of both the strong hand and the weak hand yielded a further improvement in recognition accuracy over modeling only the strong hand. A closer look at the recognition errors in the experiment modeling only the strong hand revealed that only some of them involved two-handed signs. Assuming that all of them would be corrected by the independent modeling of the two hands, the maximum possible word accuracy was been 96.5%. Under these circumstances, the 1% improvement in word accuracy is significant.

FIGURE 6. RECOGNITION RESULTS.
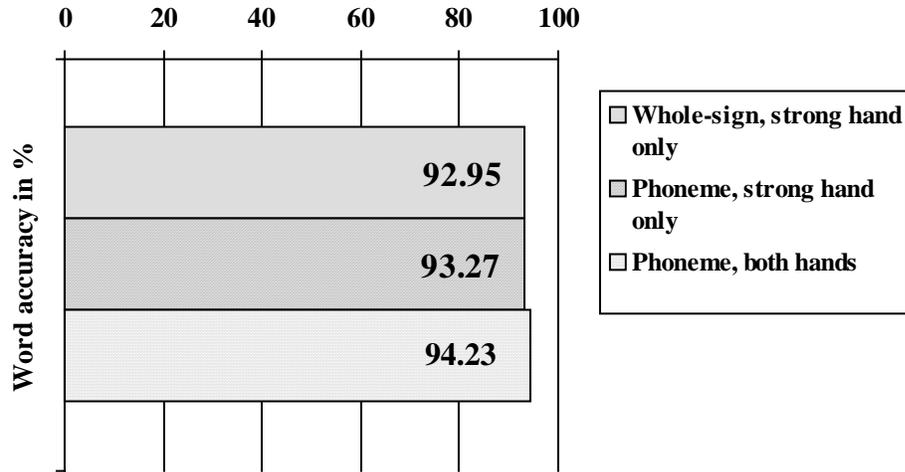
5. CONCLUSIONS.

We showed that ASL recognition can take advantage of ASL phonology. Although the Movement-Hold Model has some shortcomings that required us to make modifications to it, it gave us recognition rates comparable to whole-sign modeling and it scales much better to large vocabularies than whole-sign modeling.

We also explored the possibility of modeling the parallel aspects of ASL in independent channels. Although the independence assumption may not be fully justified from a linguistic point of view, it improved recognition rates computationally.

Future work should further explore the use of ASL linguistics in the recognition field, particularly phonology and syntax, including the use of space. It should also add other parallel aspects of sign language to the recognition framework, including the use of hand configuration, orientation, and nonmanual markers. It also needs to verify the validity of phoneme modeling and modeling of the parallel aspects with vocabularies larger than 22 signs.

REFERENCES

BRENTARI, DIANE. 1995. Sign Language Phonology: ASL. In *The Handbook of Phonological Theory*, Goldsmith, John A., Ed. Blackwell, Oxford.

COULTER, GEOFFREY. 1993. Current Issues in ASL Phonology. *Phonetics and Phonology, Volume 3*. Academic Press, Inc. San Diego, CA.

HIENZ, H. ET AL. 1999. HMM-Based continuous sign language recognition using stochastic grammars. Gesture Workshop. Gif sur Yvette, France.

LIANG, R.-H. and OUHYOUNG, M. 1998. A Real-time Continuous Gesture Recognition System for Sign Language. *Proceedings of the Third International Conference on Automatic Face and Gesture Recognition*, 558-65. Nara, Japan.

LIDDELL, SCOTT K. and JOHNSON, ROBERT E. 1989. American Sign Language: The Phonological Base. *Sign Language Studies* 64.195-277

RABINER, LAWRENCE. 1989. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*.

SANDLER, WENDY. 1989. Phonological Representation of the Sign: Linearity and Nonlinearity in American Sign Language. *Publications in Language Sciences* 32. Foris Publications. Dordrecht.

STARNER, THAD and PENTLAND, ALEX. 1995. Visual recognition of American Sign Language using Hidden Markov Models. International Workshop on Automatic Face and Gesture Recognition. Zurich, Switzerland.

STOKOE, WILLIAM C. 1960. Sign Language Structure: An Outline of the Visual Communication System of the American Deaf. *Studies in Linguistics: Occasional Papers 8*. Linstok Press. Silver Spring, MD. Revised 1978.

VOGLER, CHRISTIAN and METAXAS, DIMITRIS. 1997. Adapting Hidden Markov Models for ASL Recognition by Using Three-dimensional Computer Vision Methods. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 156-61. Orlando, FL.

VOGLER, CHRISTIAN and METAXAS, DIMITRIS. 1999a. Toward scalability in ASL recognition: breaking down signs into phonemes. Gesture Workshop. Gif sur Yvette, France.

VOGLER, CHRISTIAN and METAXAS, DIMITRIS. 1999b. Parallel Hidden Markov Models for American Sign Language Recognition. To appear in the *Proceedings of the International Conference on Computer Vision*. Kerkyra, Greece.

USING MULTIPLE MACHINE TRANSLATION PACKAGES TO PRODUCE 'AVERAGE' OUTPUT

DAN TAPPAN

Computing Research Laboratory
New Mexico State University

1. INTRODUCTION.

Machine translation (MT) is one of the oldest computer applications. For most of its half-century history, the field has remained almost exclusively within the research and intelligence communities. The few commercial products available were not intended for the individual consumer. Only recently have great advances in personal computing created a market for such products. The simultaneous explosion of the Internet provided the other integral component. Ordinary consumers now readily access information in almost any country, and especially, in almost any language. With these expanded horizons naturally comes the need for a means to read this multilingual information.

Numerous commercial MT packages are now available. As with any software, the features and capabilities differ, usually in proportion to price. What does not tend to differ, however, is the relatively low translation quality. MT is still too young and underdeveloped to produce high-quality translations of general-purpose text. While the average Internet user may not be concerned with perfect translations of web pages, other users may expect or demand better results.

This work focuses on improving translation quality. The objective is to demonstrate a proof-of-concept software architecture for manipulating translations produced simultaneously by three MT packages. This so-called *average* translation should reflect the best parts of the translations from each package. As a prototype system, this work should be considered still in progress.

2. IMPLEMENTATION.

A numerical average is defined as the sum of a series of values divided by the number of values. It serves as a *smoothing function* by leveling the highs and lows of the series to provide an overview. The average value is found somewhere inclusively between the best value and the worse value. When it is not known which values are better or worse, the average provides a safe middle ground. The use of a smoothing function can be extended to average translated sentences. The safe middle ground of the average sentence is desirable because MT packages cannot generally evaluate the quality of their translations.

Arithmetic operations are not defined over sentences, so averaging is instead handled by consensus: the average sentence is built from the words that agree most often in form and position in the set of sentences. Similarity between sentences increases the probability of correctness because each MT package independently generated the similar parts. Averaging capitalizes on the individual packages having their own strengths and

weaknesses. The intent is to exploit and augment these properties, respectively, to select those parts of the output from each package that best contribute to a composite, average sentence of overall higher quality. Furthermore, unlike a numerical average, an average sentence may be *better* than any sentence that contributed to it—the whole being greater than the sum of its parts, so to speak. These properties stem from different package implementations, as well as the extent of their resources like lexicon and grammar size, etc. The output from various packages is in many cases quite similar, but it often contains sufficient different and/or extra information to make averaging worthwhile.

The consensus approach is implemented through a principle known as *n-modular redundancy*: when several computers running similar programs on the same input data are organized in parallel, they naturally compensate for a variety of output errors. For example, if two of three computers produce the same output, the third is probably wrong. The redundancy architecture shown in Figure 1 takes as input a single source-language sentence. This sentence is simultaneously fed into MT packages *a*, *b*, and *c*, which independently translate it. The individual translations are then collected in the *averaging module*, which determines and outputs the average translation.

FIGURE 1. AVERAGING ARCHITECTURE.
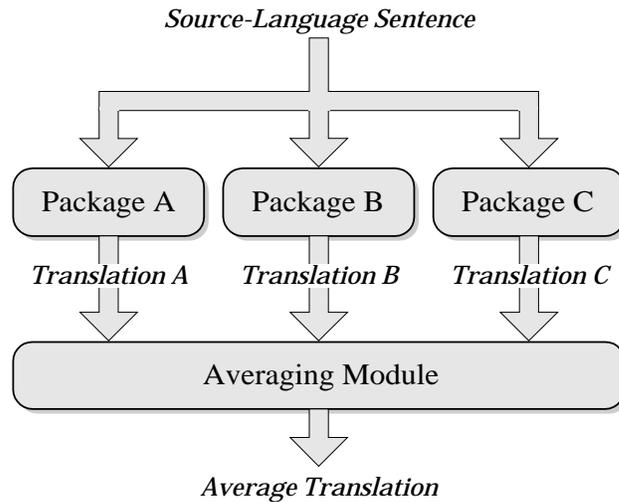
3. OPERATION.

In this work, averaging is used to compensate for three types of common translation errors:

    a.   An incorrect word appears in a correct position; e.g., sentence (2) c., where 'cat' should be 'dog'.
    b.   A correct word appears in an incorrect position; e.g., sentence (3) b.
    c.   An incorrect word appears in an incorrect position; e.g., sentence (3) b. with 'cat' instead of 'dog'.

A ranking is assigned for each word in each sentence by comparing it to the words in the other sentences. If the same word appears in the same place in all three sentences, it gets the highest ranking as both the correct word and the correct position. If the word and/or its position do not agree completely between all sentences, the ranking is reduced accordingly. The average sentence consists of the highest-ranked words in the highest-ranked positions.

Sentences (1) a.-c. illustrate the best case, where all the words in all three sentences agree in both form and position. Such convenient agreement usually occurs only for the simplest sentences.

    (1) a. *The big dog.*
          b. *The big dog.*
          c. *The big dog.*

Sentences (2) a.-c. illustrate a single error: the incorrect word 'cat' in the correct position. Sentences (2) a.-c. all agree on 'the' and 'big' as the correct words for the first and second positions, respectively. Sentences (2) a. and b. agree on 'dog' as the correct word for the third position, which outranks the lone 'cat' from (2) c. The complementary case, where a correct word appears in an incorrect position, is handled similarly.

    (2) a. *The big dog.*
          b. *The big dog*.
          c. *The big cat.*

Sentences (3) a.-c. illustrate two errors: the correct words 'dog' and 'big' in the incorrect positions, and the incorrect word 'cat' in the correct position. Again, sentences (3) a.-c. agree on 'the' as the correct word for the first position. Sentences (3) a. and c. agree on 'big' as the correct word for the second position, which is further supported by its appearance in sentence (3) b. None of the sentences agrees directly on the correct word for the third position. Indirect analysis selects 'dog' from its appearances in sentences (3) a. and b., where it outnumbers the single appearance of 'cat' in (3) c. Since the third position is the only one remaining, it is likely the correct one.

(3) a. *The big dog.*
    b. **The dog big.*
    c. *The big cat.*

Other combinations of multiple errors in word form and position are handled in a similar manner. Of course, if the sentences have too little agreement, and therefore no consensus, averaging is unlikely to find any valid solution.

4. RESULTS.

The true test of sentence averaging is how well it performs on actual output from MT packages. However, having the packages produce the test sentences makes it difficult and unwieldy to control the wide range of test parameters. Instead, for the results here, a flexible test generator was employed. For each test, the generator created a group of three sentences S_{1-3} of equal length with 1 to 15 words each and 0 to 120 total random errors between them. These sentences correspond to translations *a*, *b*, and *c* in Figure 1, respectively. A fourth sentence $S_{correct}$ of equal length was generated with no errors. As the correct sentence, its role was to evaluate the quality of S_{1-3}. It has no correspondence in Figure 1 because it was available only in the test environment; with actual translations, correctness cannot be determined automatically.

Since consensus-based averaging relies purely on comparing which words appear in which positions, the actual English grammaticality of S_{1-3} and $S_{correct}$ was not considered. Grammaticality was instead relative only to $S_{correct}$. The test generator was therefore free to create any sequence of words from its limited lexicon. Furthermore, since straightforward pattern matching is employed for comparison, the test generator and the averaging module are entirely language independent. For example, German sentences could be averaged as easily as sentences containing arbitrary symbols (e.g., 'A', 'B', etc.).

S_{1-3} were each compared against $S_{correct}$ to determine their correctness. The numerical average C_{group} of these three values reflected the correctness of the entire group. S_{1-3} were then fed through the averaging module to generate the average sentence $S_{average}$. It was also compared against $S_{correct}$ to determine its correctness $C_{average}$. The difference between $C_{average}$ and C_{group} indicated the relative improvement: a factor of two, for example, meant that $S_{average}$ was roughly twice as correct (i.e., had half as many errors).

For each of the 1,800 combinations of word length and total error count, 1,000 random tests were run, and an average improvement factor was recorded. The offset conical shape of the results in Figure 2 illustrates an obvious relationship: longer sentences with fewer errors average better than shorter sentences with more errors. This performance is understandable because having more errors reduces the likelihood of consensus between S_{1-3}. The best performance—an improvement factor between 3.5 and 4—is centered near three errors in 13-word sentences. Regardless of sentence length, no significant improvement is found after 15 errors. The worst performance reflects no improvement.

FIGURE 2. RESULTS USING THREE SIMULATED MT PACKAGES.
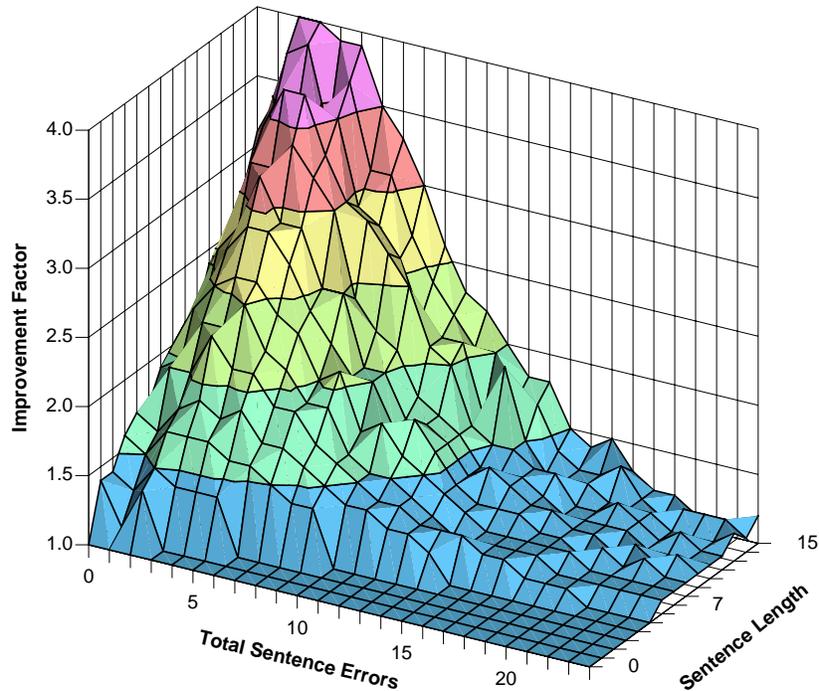
5. CONCLUSIONS.

Averaging ideal test sentences results in marked improvement in quality. This improvement depends on the error proportion between a group of sentences and their length. Longer sentences can generally tolerate more errors—up to approximately 15— before no further improvement occurs.

For this prototype architecture, two simplifying assumptions were made about the simulated MT package output: the sentences have similar, overlapping content, and they are all the same length. Unfortunately, actual output would be more difficult to average because natural language can correctly express equivalent ideas in numerous different ways. While the sentences might all have the same meaning, the particular words and ordering, as well as the sentence lengths, can vary greatly. Averaging sentences based solely on the words and their positions does not account for this limitation. Future work would need to introduce more flexibility into handling such problems. Performance would likely suffer, but it is expected that averaging would still show some improvement.

WYANDOT PHONOLOGY: RECOVERING THE SOUND SYSTEM OF AN EXTINCT LANGUAGE

CRAIG KOPRIS

State University of New York at Buffalo

1. INTRODUCTION.

Wyandot is a Northern Iroquoian language, related to Seneca, Mohawk, and Cherokee, among others. While the last speakers lived in the 1960s, the last extensive fieldwork was carried out by Marius Barbeau in 1911-1912. He intended to produce the usual trio of grammar, dictionary, and texts, but did not finish. No grammar was started, and the dictionary has remained in manuscript form, probably due to the extensive misanalyses present. Only the texts were ever published, in 1960. These consisted of approximately 253 pages of legible handwritten transcription of Wyandot and typed word glosses, with an additional 51 typed pages of free translation.

To perform a phonological analysis of a language, linguists usually phonetically transcribe utterances obtained from native speakers, then search the data for contrastive and complementary distribution, seek minimal pairs, return to the speakers for verification of hypotheses, and so on. But what should be done when there are no longer any speakers of the language? How does one proceed when the linguist cannot simply go to a speaker to round out a paradigm, or run a spectrogram? This is the problem in sorting out the phonology of Wyandot.

2. WRITTEN TRANSCRIPTION.

Since there is a lack of fluent native speakers of Wyandot and, therefore, the texts cannot be retranscribed phonetically, the transcription system used by Barbeau himself will be the basis used here instead. This particular system has several drawbacks, to be discussed in the following sections. Some peculiarities can be assigned to the time period when the fieldwork was done, centering around 1912. However, this does not explain the retention of the system in publishing Barbeau (1960). Our analysis will combine standard phonological procedure, the comparative method, and historical insights. Various types of brackets will be used for different transcriptions: Barbeau's transcription will be in <angle brackets>, the phonetic realization in [square brackets], and the phonemic interpretation in /slashes/.

2.1. Barbeau's Characters. Barbeau gives a description of his transcription system in several works, with minor variations. The following description closely follows the format Barbeau used, with two exceptions. Additional notes found only in certain works are added

below the standard description, and modern IPA equivalents are placed to the side.

VOWELS

a   as in English m*a*t, and in French p*ara*de   [æ ~ a]

    Barbeau (1915a) hedges with "a vowel closely resembling those in"

e   as in French *é* - English *a* in c*a*ve   [e ~ e^j]

ɛ   as in French *è* - English *e* in p*e*t   [ɛ]

i   as in French *i* - English *i* in f*i*t   [i ~ ɪ]

u   as in French *ou*, - English *o* in l*o*se   [u ~ u^w]

NASALIZED VOWELS

ą, ɛ̨, į̇ (rare), the *a, e, i* above vowels nasalized:

ą   as in the French march*an*d   [ɑ̃]

ɛ̨   as in the French *in*, in v*in*   [ɛ̃]

ǫ   the open *o* nasalized as in French b*on*   [ɔ̃]

SEMIVOWELS

w   as in the English *wine*   [w]

y   as in the English *yes*   [j]

CONSONANTS

c   as in the English *she*, - French *chat*   [ʃ]

j   as in the French *jamais* followed by a brief *y*   [ʒ^j]

    Barbeau (1915a) hedges, with "closely resembling that of French *jamais*" and adding that the

    y is very brief

s   as in the English *sit*   [s]

t   as *t* with a slight aspiration   [t̪ ~ t^h]

    Barbeau (1915a) states "approximately as in English and French"

    Barbeau (1949) also indicates "followed by a slight aspiration"

d   as in *done*, often preceded by a weak *^{n}*   [t ~ d ~ ^{n}d]

Barbeau (1915a) hedges with "approximately"

Barbeau (1915b) refers to a "weak *n*"

k   as in *key*   [k ~ k^h ~ c ~ c^h]

Barbeau (1915a) states "approximately as in English"

g   the sonant *g* followed by a *y*, often preceded by a weak ^{n} [gj ~ ^{ŋ}gj ~ ɟ ~ ^{ɲ}ɟ]

Barbeau (1915a) uses *gy*, with a "preceding weak *ñ*"

Barbeau (1915b) also uses *gy*, described as "sonant *g* immediately followed by *y*, often with a preceding weak ^{ŋ}, palatalized *ng* of English *sing*"

ķ   *k* followed by *y*   [k^j ~ k^{hj} ~ c ~ c^h]

Barbeau (1915a,b) instead use ky

m   and   [m]

n   as in English and French   [n ~ n̪]

ñ   as the *ñ* in Spanish; the *gn* in Italian   [ɲ]

r   corresponding to the English *r*   [ɹ]

Barbeau (1915b) hedges with "roughly"

ŗ   rare; deep palatal, tending to disappear   ?

Barbeau (1915a, b) lack this character

h   aspiration always followed by a vowel   [h]

Superscript letters indicate very brief, and sometimes unvoiced consonants and vowels, as in

^{n}dătrá^ʼaskwĭjù^{ʼun}di^ʼ

DIACRITICAL MARKS

ʼ   glottal stop or catch as in ^{n}gá^{ʼa}wic   [ʔ]

ʻ   breathing after a vowel and before a consonant as in aʻcɛʻk   [h]

´   over a vowel shows the main stress or accent in a word; it usually corresponds to a rising pitch of the voice

Barbeau (1915b) refers to this as "high pitch"

ˋ   minor or weaker accent

Barbeau (1915a) uses "secondary"

·   a raised period after a vowel indicates that it is long, as in ĭyǫ́·tɛˀ

˘   over a vowel makes it brief: tĕhát

Unmarked vowels are of medium length. Two consecutive brief vowels may be combined into one main accent: ŭrḝhą̆ˀ (N.B. the two ˘ are linked above by ⌢, with a ´ above the link [CK])

Difficulties in ascertaining the phonetic nature of the Barbeau characters are readily apparent. For instance, vowels are often described as like equivalents in French and English, even though French and English vowels are often quite different phonetically. The offglides on English [eʲ uʷ] do not appear on French [e u], for one example. Also, although English-speaking beginning students of French may often perceive French [a] as [æ], the two are distinct. Although the nasal vowels are described as nasalized versions of the oral vowels, the descriptions of oral <a> as [æ ~ a] do not correspond to the nasal <ą> as [ɑ̃].

At times Barbeau describes <t k> as "unaspirated surds", while at other times <t> is contradictorily described as having a "slight aspiration". Further, <t n> described as similar to English and French obscures the actual place of articulation, which in English is alveolar but in French is dental. The "deep palatal" <ɹ̣> is not given an IPA equivalent, since it is not clear what Barbeau meant by the description.

Additional characters not mentioned but used include:

ᵐ   in Barbeau (1915a), appearing in the environment uᵐw

ñ   in Barbeau (1915a), appearing in the environment ñgy

ŋ   in Barbeau (1915b), appearing in the environment ⁿgy

ç̧   rarely used, in Barbeau (1960)

ʷ   rarely used, in Barbeau (1960)

ʸ   rarely used, in Barbeau (1960)

The characters <ᵐ ñ ŋ> are apparently <ⁿ> assimilating to a labial or velar environment. On the analogy k:ḵ::c:c̨, <c̨> is presumably equivalent to <cy>. The "superior letters" refer primarily to <ⁿ>, which usually appears before d or after a nasal vowel, as well as to echoed vowels after glottal stops, i.e., ⁱ ᵉ ḝ ᵋ ᶓ ᵃ ą̆ ᵒ ǫ ᵘ.

Some of the variations can be explained according to Barbeau's personal notes, *inter alia*, as character alternations based on typesetting demands. For typesetters unable to create <ḵ>, Barbeau would use <ky>, for those unable to print <ɛ́> he would use <ɛ´>, and so on.

3. OVER-DIFFERENTIATION.

In addition to the previously described peculiarities, the Barbeau system has several other drawbacks. One of these is over-differentiation. For instance, the same word may appear with several different length, stress, and nasalization patterns that do not contrast, as in (1):[1]

(1) a. ăh̀ɛ̀hǎǫ́ˀ *he said* TN:02:068:03
   b. ah̢ɛ̀haǫ́ˀ *he said* TN:12:113:18
   c. ah̢ɛ̀hǎǫ́ˀ *he said* TN:01:062:20; TN:02:063:36
   d. ăh̢ɛ̀haǫ́ˀ *he said* TN:12:114:15
   e. ăh̢ɛ̀hǎǫ́ˀ *he said* TN:01:060:09
   f. ah̢ɛ̀hǫ̆ǫ́ˀ *he said* TN:12:114:58
   g. ah̢ɛ̀ ⁿhǎǫ́ˀ *he said* TN:01:060:03

Each of these examples shows a different pattern, as demonstrated in (2):

(2) a.
   b.
   c.
   d.
   e.
   f.
   g.

Note that the only vowel that is consistent is the last. All of the others vary in length, stress, and nasalization.

 Vowels may also appear with different qualities, which do not contrast, as can be seen in (3):

(3) a. aˀᵃwé·hǫˀ *she thought* TN:04:082:24
   b. aˀᵃwɛ́·hǫˀ *she thought* TN:01:059:14; TN:04:079:41

However, these two sounds cannot simply be considered allophones, since <ɛ> enters into alternations that <e> does not. <ɛ> alternates with <ɛ̧> as in (4); however, <e> and <ɛ̧ not alternate:

(4) a. yăñɛ́·nǫ‘ *dog* TN:16:127:62
   b. yăñɛ̧́nǫ‘ *dog* TN:16:126:36

This time there is a noncontrastive difference between <ɛ> and <ɛ̣>.

Any phonemicization of Barbeau's materials must take into consideration that not all of his characters represent distinct phonemes, but rather may indicate minor variations in sound.

3.1. Under-differentiation. Another problem with the Barbeau orthography is that it is under-differentiated. That is, not all distinctive contrasts are marked.

All other Iroquoian languages contrast the simple phonemes /t/ and /k/ with the clusters /th/ and /kh/. Example (5) shows two forms based on the Cayuga root -ehst- *use* (Mithun and Henry 1982). (Underlining signifies devoicing in Cayuga orthography.)

(5)   a.   tesẹn?ǫá-ehst-ǫh *you are using your head*
        b.   s-e̱hst-ha? *you use it*

The first form adjoins t and ǫ across a morpheme boundary, while the second has t and h. These are usually phonetically realized as [t k] and [tʰ kʰ]. They are frequently represented as <d g> and <t k> in orthographies developed by English speakers or intended for English-speaking learners.

Neither stop-h cluster appears in the Barbeau texts. Either the transcription system fails to represent the distinction between /t k/ and /th kh/, or Wyandot is the only Iroquoian language to have lost the distinction. Knowing that <d g> both appear in the texts, it may be hypothesized that the Barbeau orthography is English-based. However, <d> is cognate with /n/, not /t/, in the other languages. Compare in (6) forms based on the root for *drum*/*barrel* in Wyandot (6a) and Seneca (6b) (Chafe 1967):

(6)   a.   hụ̀ⁿ-ʼⁿdaʿkw-ă̓éhă̱ke̓ *he the drum goes on beating* TN:03:076:09
        b.   ye-?nǫhkw-a?éstha? *drumstick*

Here the initial ? of the morpheme as well as the final hkw, are the same in both languages. The ǫ in Seneca comes from the rule in (7):

(7)   a → ǫ / n _

This leaves Seneca /n/ cognate with Wyandot <ⁿd>. Therefore, Barbeau did not interpret unaspirated [t] as an English speaker would, as /d/, and thus did not similarly misinterpret an aspirated [th] cluster in English terms, as /t/.

Simple English-biased misrepresentations of [kh] and [k] in a parallel manner as [k] and [g] are not the proper explanation either, as [g] is an allophone of /d/ (see 4.1).

Furthermore, sometimes Wyandot <t> corresponds to /t/ in cognate languages, and sometimes to /th/. In (8) compare the root <hati> *they* in Wyandot (8a) and -hati- *they* in Cayuga (8b):

(8) a. ǎ‒hàti‒ˑcrǫ́ⁿgaʼ *they make* TN:07:100:40
    b. hati‒ʔtreḥtǫ́ːnihs *they make cars*

Here the Wyandot stop is equivalent to a Cayuga simple stop. But sometimes Wyandot <t> has /th/ correspondences. Compare <tɛʼt> *pound* in Wyandot (9a) with -theʔt- *grind; pound* in Cayuga (9b):

(9) a. sǎíˑ‒tɛʼt *again they (2) pound corn* TN:04:083:09
    b. akate-theʔt-hróʔktʔaǫh *I am out of flour*

or Wyandot <ratɛ> (10a) and Cayuga -rathę- (10b) *climb*:

(10) a. hǎ‒rátɛ‒ʼ *he climbs* TN:13:118:20
     b. haʔt-ráːthę-h *climb up there!*

In each of these cases Wyandot <t> corresponds to Cayuga /th/. The same can be shown for Wyandot <k> and /k kh/ cognates in Mohawk. Example (11) compares Wyandot <gak> *marry* (11a) and Mohawk -nyak- *get married* (11b) (Mohawk from Michelson 1973).

(11) a. ahǎhu‒gák‒aʼ *he got married* TN:18:134:29
     b. waké-nyak-s *I get married*

Note that here Wyandot <k> is equivalent to Mohawk /k/. In the following example, Wyandot <tsiʼkɛʼt> (12a) is contrasted with Mohawk -tsikheʔt- (12b) *sugar*.

(12) a. du‒tsíʼkɛ̀ʼᵋt‒aʼ *in a sugar lump* TN:14:123:20
     b. o-tsikhèːt-a *sugar*

As can be seen, this equivalency is between <k> and /kh/. Thus, simple stops in Barbeau's orthography sometimes represent simple stops, and sometimes represent aspirated clusters, based on comparison to related languages.

Thus any phonemicization of Barbeau's materials must take into consideration that not all of his characters represent individual phonemes, but rather may indicate clusters of distinct phonemes.

4. CONSONANTS.

The various consonants Barbeau uses are: <t ḵ k ꞌ ⁿt ⁿḵ ⁿk d g ⁿd ⁿg s c ç h ꜥ j m ᵐn ⁿ ᵑ ñ ñ r r̥ w ʷ y ʸ >.

Since <ḵ ç ñ> represent clusters, they can be eliminated in favor of those corresponding clusters: <ky cy ny>. The description of <g> always refers to an offglide, so this is better represented as <gy>. This reduces the inventory to <t k ꞌ ⁿt ⁿk d g ⁿd ⁿg s c h ꜥ j m ᵐn ⁿ ᵑ r r̥ w ʷ y ʸ >.

4.1. Allophones. Further reduction can be obtained by eliminating allophones. The voiced stops are in free variation with prenasalized voiced stops. Thus, <d> and <ⁿd> are in free variation, as in (13), and <g> and <ⁿg> are, as in (14).

(13) a. tayĕmȩ̀·‑d‑ű̆rḝha̧ꞌ *thou my desire findest out* TN:04:088:18[2]
      b. ăhămȩ̀·‑ⁿd‑ű̆rḝha̧ꞌ *he her desire finds out* TN:04:088:47

(14) a. ăhì·‑g‑ắha *they (2) eat* TN:04:081:34
      b. ăhăti‑ⁿg‑ắha· *(for) them to eat* TN:03:076:28

Since <dy> does not occur, and <g> is described as followed by *y*, we can eliminate <g> through the rule <d> → <g> / _ [y]. Furthermore, as can be seen in (15-16), <g> is cognate with /ni ~ ny/ elsewhere in Iroquoian.

(15) a. ăhá‑ꜥr‑ǒ̆ⁿg‑aꞌ *he makes* TN:01:062:06
      b. ase‑hsr‑ǫ́ːni‑ʔ *you made it*
      c. k‑ǘːni *I make*

(16) a. ăñǫ̀·‑gák‑a *I you married* TN:02:068:14
      b. aːyǫkhnı́ː‑nyaːk *we two would marry*
      c. waké‑nyak‑s *I get married*

The root for *make* ends in ‑ni‑ in Cayuga (15b) and Mohawk (15c), but <g> in Wyandot (15a). *Marry* begins with ‑ny‑ in Cayuga (16b) and Mohawk (16c) but <g> in Wyandot (16a).

Most importantly, roots can be found within Wyandot which show <d ~ g> alternation:

(17)   a.   yà·−ré·ⁿd−a³ *trap* TN:05:091:29
       b.   hă−rɛ̨ⁿd−ărǫ́·ñǫˤs *he his traps sets several* TN:16:126:23
       c.   ahăyắ−rɛ̨ⁿg−a³ *he trap takes away* TN:05:091:24

In (17) the root for *trap*, <rɛ̨ⁿd>, varies between final <d> in (17a-b) and final <g> in (17c).
    The prenasalized voiceless stops are in variation with the simple versions, appearing after a previous nasal segment.

(18)   a.   ăhu·nǫ́−ⁿt *she (to) him gave* TN:15:125:58
       b.   ɛ̆yǫ·nǫ́−t *will I (to) you give* TN:15:125:43

Here two forms of *give* show free variation of <ⁿt> and <t> after a nasal <ǫ>. In (19) <ⁿk> varies with <k> after a nasal.

(19)   a.   da³yă̆ɛ̨−ⁿk−wa³ *that she planted (the seeds)* TN:01:061:11
       b.   hĭñɛ̨−k−wa³ *they two are to plant* TN:04:090:30

    A rule of the form <r> → <r̥> / _ <h> eliminates the "deep palatal":

(20)   a.   dă̆³ú·−r̥h−ɛhą³ *the next day* TN:04:088:05
       b.   dĭyá·−r̥h−i³ *that tree* TN:03:074:28

    Almost all instances of <m> can be derived from a <w> in the environment of a nasal segment. <w> does not appear adjacent to a nasal, and <m> does not appear without another nasal. This can be covered by the rule <w> → <m> / (³) % Ɣ. (Where % indicates either a preceding or following environment.)

(21)   a.   ăhă̆−mɛ̨−·ⁿdǘrɛ̨hą³ *he her desire finds out* TN:04:088:47
       b.   ăhàtĭkaskà³ᵃh−ǫ́m−ą³ *they a division made* TN:07:098:48
       c.   ahăyǫñ−ǫ̀³m−ąˤcɛ̨ⁿdré· *he their bills tied up* TN:11:110:16

However, a few unexplained instances have <m> adjacent to <ą>, which itself is not phonemic (see 5.1).

(22)   mą̆−hąˤcănǫ́· *most so* TN:02:064:05

These examples are almost all particles, however, and it is known that particles can have phonological peculiarities (Mithun 1982).

Since <ʿ> is "breathing after a vowel and before a consonant" and <h> is "aspiration always followed by a vowel", it is clear that they are allographs, with <h> used where English has /h/, and <ʿ> used where English does not allow /h/.

As mentioned previously, $<^{m\,\tilde{n}\,\eta}>$ are assimilations of $<^{n}>$ to labial and velar environments. $<^{n}>$ itself almost only occurs in the prenasalized stops, which have already been dealt with. The exceptions occur after nasal vowels, as in (23):

(23) a. ah−ξ̀ⁿ−hǎǫ̀ʾ *he said* TN:01:060:03
      b. kuʾᵘg−ξ́ⁿ−seʿ *snake* TN:08:102:18

These allophones, allographs, and other variations reduce the set of consonants to the following: <t k ʾ d s c h j n r w y>.

5. VOWELS.

Barbeau uses 62 vowel characters, plus an additional 10 raised characters: $<^{i\ e\ ę\ \varepsilon\ \xi\ a}$ $^{ą\ o\ ǫ\ u}>$. The 62 are shown below:

i	í	ì	ĭ	í̋	ì̏		í̧	ì̧
e	é	è	ĕ	é̋	è̏	ę	ę́	ę̀	ę̆	ę́̋	ę̀̏
ε	έ	ὲ	ε̆	έ̋	ὲ̏	ξ	ξ́	ξ̀	ξ̆	ξ́̋
a	á	à	ă	á̋	à̏	ą	ą́	ą̀	ą̆	ą́̋
o	ó	ò	ŏ	ó̋	ò̏	ǫ	ǫ́	ǫ̀	ǫ̆	ǫ́̋	ǫ̀̏
u	ú	ù	ŭ	ú̋		ų	ų́	ų̀

The diacritic patterns can be shown in matrix format for clarity.

plain	´	`	˘	ˊ	ˋ	´̦	`̦	˘̦	ˊ̦	ˋ̦
i	+	+	+	+	+	+	-	+	+	-	-	-
e	+	+	+	+	+	+	+	+	+	+	+	+
ε	+	+	+	+	+	+	+	+	+	+	+	-
a	+	+	+	+	+	+	+	+	+	+	+	-
o	+	+	+	+	+	+	+	+	+	+	+	+
u	+	+	+	+	+	-	+	+	+	-	-	-

5.1. Allophones. These 72 characters can be reduced by eliminating allophones. <e> and <ε> alternate optionally.

(24)   a.   ăhàti‑jé·rat *they used* TN:07:099:05
        b.   àhă‑ɛra’t *he used* TN:13:119:31

In (24) the root <(y)era’t> *use* alternates between <e> and <ε>.
In a parallel fashion, the nasal equivalents <ę> and <ɛ̨> also optionally alternate.

(25)   a.   ăhatïhà‘c‑ę́‑ta’ *they a council held* TN:07:098:42
        b.   hătĕhá‘c‑ɛ̨‑)ᵋ *they a council hold* TN:09:104:32

This shows a vowel quality alternation in the root <ha‘cęt> *hold council*.
Unfortunately, there is also a nasal-oral alternation between <ɛ̨> and <ε>, though this is in the vicinity of another nasal segment.

(26)   a.   a’tĕm‑ɛ̨‑taye‘ *every day* TN:01:059:03
        b.   a’tĕm‑ε‑ⁿtá·ye‘ *every day* TN:12:112:23

In (26) <ɛ̨> and <ε> are in free variation after nasal <m>. Some or all of these may be due to errors in transcription. That is, the nasal nature of a nearby nasal segment might have masked nasalization on the vowel itself. These variations may thus actually be English-biased perception mistakes, which would also account for the several other nasal-oral vowel alternations <i ~ į>, <a ~ ą>, <u ~ ų>, and <o ~ ǫ>.
The character <į> only appears next to a nasal, and <i> does not appear next to nasals. Thus, the alternation of <i> and <į> is the result of nasalization from another nasal segment.

(27) a.   d‒ǫ̆ì‒ⁿdăwé²ᵉt *we two are brother and sister in law* TN:04:084:14

      b.   n‒ǫ̆i‒ⁿdăwé²ᵉt *we two are brothers and sister in law* TN:04:084:23

Here the alternation of \<i̧\> with \<i\> can be explained through the presence of the preceding nasal \<ǫ\>.

     The same holds for \<ą\>, except that a laryngeal \<h ˁ ˀ\>can intervene.

(28) a.   ăh‒ą̀·ⁿg‒á̧ˁ *he her chases (after)* TN:02:069:26

      b.   ăhà tı̆kaskà²ᵃh‒ǫ̆́mą‒² *they a division made* TN:07:098:48

      c.   ahăjá²ᵃtŭr‒ę̀·hą‒² *he me finds* TN:01:061:27

Example (28a) shows \<ą\> before a nasal \<ⁿg\> while (28b) shows \<ą\> after nasals \<ǫm\>. (28c) shows nasal assimilation across an intervening larygeal \<h\>.

     As mentioned previously, sometimes \<ą\> appears without a nasal environment, especially in particles.

(29)   hą²ᵃrą² *only* TN:04:081:37; TN:04:087:48; TN:04:088:14

\<u\> can appear as nasalized \<u̧\> in the vicinity of a nasal, in that oral \<u\> does not occur next to a nasal, and \<u̧\> only appears adjacent to a nasal. Example (30) shows the occurrence of \<u̧\> adjacent to a nasal.

(30)   t‒ú̧·ⁿd‒i *also* TN:02:065:27

The opposite occurs with \<ǫ\> and \<o\>, where \<o\> only appears adjacent to a nasal. That is, where \<ą u̧\> are nasal allophones of \<a u\>, \<o\> is an allophone of \<ǫ\>.

(31) a.   kasăḵá²ḵe‒nǫ *here thou liest down* TN:04:086:09-10

      b.   ăhą̆̀ḵa²ḵĕ‒no *he lies down* TN:12:112:42

In (31) we see an alternation of \<ǫ\> and \<o\> following \<n\>. As with \<ɛ\> as allophone of \<ę\>, this may be a transcription error, the nasality of a nearby nasal segment hiding the nasality of the vowel \<ǫ\>.

Finally, vowels may be echoed optionally after a glottal stop.

(32) a. yăhà·w‑ȋⁱ *she carries* TN:04:090:26
      b. yǫ̀ⁿtăr‑é⁾ᵉ *lake* TN:08:103:50
      c. ăhǻt‑ɛ·⁾ᵋ‑wa⁾ *he runs off* TN:11:109:13
      d. y‑ɛ́⁾ᵋ‑ric *lion* TN:09:105:33
      e. ⁾eˤskĕmɛⁿdăr‑ǽ⁾ᵃ‑cǫñǫ⁾ *will thou (with) me converse* TN:04:079:10
      f. ŭwa⁾h‑ǽ⁾ᵃ‑rą⁾ *another one* TN:03:075:08
      g. c‑ǫ⁾°‑mǻ *by far* TN:04:080:49
      h. yɛˤtĳ‑ú⁾ᵘ‑rŭn‑ǫ⁾ǫ *prairie turtle tribe* TN:09:105:13

This example shows each of the echoed vowels appearing after a glottal stop. The nasal-oral alternations discussed above also occur amongst the echoed vowels, as can be seen in (32g), where oral <°> appears with nasal <ǫ>.

6. PHONEMES.

With the various reductions and alternations in sections 4 and 5, and modernizing the symbols to a traditional Americanist set, we can obtain the phonemic inventory of Wyandot. The consonant phonemes are listed below:

(33) Consonant Phonemes

	labial	alveolar	palatal	velar	glottal
voiceless stop		t		k	ʔ
voiced stop		d			
voiceless fricative		s	š		h
voiced fricative			ž		
nasal		n			
resonant		r			
glide	w		y		

The term "alveolar" is used here with the caveat that Barbeau was not explicit in his description of the place of articulation, as he described it as similar to both English (and thus alveolar), and to French (and thus dental).

There is a single voiced stop, /d/, contrasting with a voiceless stop /t/. This is interesting in that no other Iroquoian language has a voicing distinction. Furthermore, another unusual feature is the presence of a voiced fricative /ž/. This is again counter to Iroquoian languages'

lack of a phonemic voicing distinction. It is also unusual cross-linguistically, with a /š ž/ contrast without a corresponding /s *z/ contrast.

Phonotactics are more problematic, in that the aspirated clusters /tʰ kʰ/ are unmarked in the Barbeau orthography. However, we can assess consonant clusters aside from these. The initial consonant clusters (#CC) include: ts‑, tr‑, kw‑, ky‑, dr‑, dy‑, st‑, sk‑, šr‑, šy‑ and ny‑, as shown in the following table:

(34) #CC Clusters

	t	k	s	r	w	y
t	-	-	+	+	-	-
k	-	-	-	-	+	+
d	-	-	-	+	-	+
s	+	+	-	-	-	-
š	-	-	-	+	-	+
n	-	-	-	-	-	+

There is a single #CCC cluster, skw‑.

Final clusters (CC#) can consist of only two members: ‑ʔt, ‑ʔk, ‑ʔs, ‑st, ‑ht, ‑hk, ‑hs, ‑hš, ‑yh, and ‑rh, as shown:

(35) CC# Clusters

	t	k	s	š	h
ʔ	+	+	+	-	-
s	+	-	-	-	-
h	+	+	+	+	-
y	-	-	-	-	+
r	-	-	-	-	+

Medial clusters (-CC-) can of course be longer. The medial CC clusters include: ‑tr‑, ‑kw‑, ‑ky‑, ‑ʔt‑, ‑ʔk‑, ‑ʔd‑, ‑ʔs‑, ‑ʔh‑, ‑ʔž‑, ‑ʔr‑, ‑ʔw‑, ‑ʔy‑, ‑dr‑, ‑dy‑, ‑st‑, ‑sk‑, ‑sh‑, ‑šk‑, ‑šr‑, ‑ht‑, ‑hk‑, ‑hs‑, ‑hš‑, ‑hr‑, ‑ny‑, ‑rh‑, ‑yʔ‑, and ‑yr‑

(36)   -CC- Clusters

	t	k	ʔ	d	s	š	h	ž	r	w	y
t	-	-	-	-	-	-	-	-	+	-	-
k	-	-	-	-	-	-	-	-	-	+	+
ʔ	+	+	-	+	+	-	+	+	+	+	+
d	-	-	-	-	-	-	-	-	+	-	+
s	+	+	-	-	-	-	+	-	-	-	-
š	-	+	-	-	-	-	-	-	+	-	-
h	+	+	-	-	+	+	-	-	+	-	-
n	-	-	-	-	-	-	-	-	-	-	+
r	-	-	-	-	-	-	+	-	-	-	-
y	-	-	+	-	-	-	-	-	+	-	-

Medial CCC clusters are: -ʔts-, -ʔtr-, -ʔkw-, -ʔky-, -ʔdr-, -ʔdy-, -ʔsk-, -ʔny-, -skw-, -sky-, -hts-, -hkw-, -hky-, -hst-, -hsk-, -hšt-, -hšr-, -hšy-, and -hny-. Medial CCCC clusters are: -ʔskw-, -hstr-, -hskw-, and -hsky-.

The Wyandot vowel phoneme inventory can likewise be listed.

(37)   Vowel Phonemes

            i   u
             e
               ę ǫ
                 a

This can be further rearranged if desired into a symmetrical shape:

(38)   Vowel Phonemes

            i u
            ę ǫ
            e a

Although Barbeau described the nasal phonemes as /ę̃ ǫ̃/, for simplicity they can be written /ę ǫ/. Few vowel clusters can be found. These are: ae-, ai-, aǫ-, eu- initially; -VeV- medially; -ae, -aǫ, and -eu finally.

On a final note, there are various rules whose synchronic or diachronic nature cannot be determined. Some dialects of various Iroquoian languages have a rule t → k / _ y. It appears this rule also is at work in Wyandot, since there are occurrences of ky in Wyandot but no *ty. However, most examples are not adequate to decide whether the rule is synchronic or diachronic. Fortunately, there are a small number of examples of alternations between t and k before y.

(39)   a.   yărǫ́·ta꞉ *log* TN:11:109:36
       b.   ahărǫ́ⁿk̲a꞉ᵃ *he log down* TN:19:143:47

This example shows an alternation in the root -rǫt- *log* between final /k/ before /y/ (shown by the <k̲>), and final /t/ elsewhere.

(40)   a.   hŭnę́rò·ti꞉ *he goes out hunting* TN:04:087:22
       b.   ahŭnĕrǫ́k̲e꞉s *he goes out hunting* TN:12:112:19

The root for *hunt*, -nęrǫt-, appears with final /t/ except in (40b), where it appears as /k/ before the following /y/.

   Thus, there is a phonemic rule changing a voiceless stop from alveolar to velar before y, and a phonetic rule changing a voiced stop from alveolar to velar before y. The phonemic rule appears to have been extended in Wyandot to cover /w/ in addition to /y/, that is, to both glides. This can be seen comparing -kahkw- *see* in Wyandot (41a) with -kaht-hw- *look at, see* in Seneca (41b):

(41)   a.   ăhǻ–kaʿkw–a꞉ *he saw* TN:12:113:01
       b.   akát-kathw-ęh *I've looked at it*[3]

The Seneca form shows /thw/, while the Wyandot has /kw/. The cluster /tw/ does not occur in Wyandot. And the only example of <gw> is cognate with /kw/ in the other languages. There is not enough evidence to declare the rule synchronic. Thus, we have three very similar rules:

(42)   a.   t → k / _ y
       b.   /d/ → [g] / _ y
       c.   (*)t > k / _ w

These could be written as one if we mixed phonetic and phonemic notation, and diachronic and synchronic changes.

7. CONCLUSION.

A language without fluent speakers can still be analyzed phonologically, provided there are several favorable factors. One is that texts exist in a reasonably detailed, even if inadequate, orthography. Over-differentiation is easier to undo than under-differentiation, though a certain amount of under-differentiation is acceptable if there are related languages available for comparison. By using standard phonemic procedures on the written material available, and applying this data through the comparative method, one can be reasonably sure as to the extinct language's phonological system. The historical, comparative, and phonological procedures are most readily used simultaneously, however, mixing levels rather than rigorously separating them.

REFERENCES

BARBEAU, MARIUS. 1915a. Classification of Iroquoian Radicals with Subjective Pronominal Prefixes. Canada. *Department of Mines Geological Survey Memoir 46, Anthropological Series 7.* Ottawa: Government Printing Bureau.

_____. 1915b. Huron and Wyandot Mythology: With an Appendix Containing Earlier Published Records. *Canada Department of Mines Geological Survey Memoir 80, Anthropological Series 11*. Ottawa: Government Printing Bureau.

_____. 1949. How the Huron-Wyandot Language was Saved from Oblivion. *Proceedings of the American Philosophical Society* 93.226-32.

_____. 1960. Huron-Wyandot Traditional Narratives: In Translations and Native Texts. *National Museum of Canada Bulletin 165, Anthropological Series 47*. Ottawa: The Queen's Printer and Controller of Stationary.

CHAFE, WALLACE. 1967. Seneca Morphology and Dictionary. *Smithsonian Contributions to Anthropology 4*. Washington: Smithsonian Institution Press.

MICHELSON, GUNTHER. 1973. A Thousand Words of Mohawk. *Mercury Series Ethnology Division Papers 5*.

MITHUN, MARIANNE. 1982. The Synchronic and Diachronic Behavior of Plops, Squeaks, Croaks, Sighs, and Moans. *International Journal of American Linguistics* 48.49-58.

MITHUN, MARIANNE and REGINALD HENRY. 1982. Watęwayęstanih: A Cayuga Teaching Grammar. Brantford, Ontario: Woodland Indian Cultural Educational Centre.

1. The numbers following the examples are source codes referring to *Traditional Narratives*: text number: page number: word number.

2. The hyphens in these examples are separating out the phones in question and are not to be construed as morphemes.

3. The missing /h/ in the Seneca is due to a rule simplifying clusters.

AN ANALYSIS OF PASSIVE CONSTRUCTIONS IN THAI[*]

UNCHALEE SINGNOI

University of Oregon

1. INTRODUCTION.

In this paper we attempt to provide an analysis of Thai passive constructions and their grammatical relations and argue that there is a true passive in Thai. First we will present the facts that are syntactically problematic in the analysis and provide argument for Thai passives. Next we will examine different types of passives and provide both semantic and syntactic accounts for each type and seek the interaction between its semantics and syntax. For some interests, possessive raising in passive constructions will be discussed next. Then, the subjecthood in the passive constructions will be examined for their subject properties in Thai passives. Moreover, passive constructions will be synchronically distinguished from other similar clause-internal structures in Thai, say, topicalization. Finally, the text distribution of actives and passives is provided.

Before we discuss the passive constructions in Thai, we need to have background on some relevant Thai syntactic structures. Thai is described as a rigid SVO language. The main verbs and their arguments, such as subjects and objects, are mostly indicated by position. Thai is also a non-inflectional language: there are no case markers for subjects and objects; nor is there any agreement (in person, number or gender) between a verb and its subject, or between a verb and its object; nor is there any change of verb form according to the tense and time of an event or an action (Sookgasem, 1992). Although verbs are not inflected, they are accompanied by adverbs or auxiliaries which indicate tenses and aspects. To discuss passivization and grammatical relations in Thai, therefore, we do not involve the overt coding properties like verb agreement and nominal case morphology. Only word-order is a major concern.

Another important issue concerning Thai passive constructions is serial verb constructions. Since Thai has certain verbs which also function as auxiliaries or adverbs indicating tense and manner after main verbs, the co-occurrence of them with main verbs consequently results in serial verb constructions. For example, consider the clauses where the verb *chaa* 'slow' indicates the manner of the main verb *duan* 'walk' in (1), *set 'finish'* indicates the past tense of *phuut* 'speak' in (2), and *yuu 'stay'* indicates the progressive tense of *tham* 'make' in (3) as shown below :

 (1) khaw duan chaa
   3p walk slow
   'S/he walks slowly.'

(2)   chan phuut set leaw
        1p speak finish already
        'I have already finished speaking.'

(3)   mee tham kapkhaaw yuu
        mother make food live
        (Lit : Mother is making food.)
        'Mother is cooking.'

2. ON THE PROBLEMATIC ASPECTS OF A PASSIVE CONSTRUCTION IN THAI.

Before we enter the discussion of the analysis of passives in Thai, we need to be clear about the notion of "passive" and some problematic facts when the term is applied in Thai. Cross-linguistically, the application of the term "passive" has not been strictly constrained, and the term "passive" has often been used for a diverse range of structural arrangements and functions (Lakayama, 1997). The syntactic properties of "prototypical passive", as given by some scholars (e.g., Givon 1990b and Keenan 1985), can be summarized as:

(4) DEFINITION OF THE PROTOTYPICAL PASSIVE:

    a.   The passive clause structurally alternates with the active clause.
    b.   The non-agent (active object) is fully promoted to subjecthood.
    c.   The agent (active subject) is demoted via deletion or oblique
          case-marking.
    d.   The passivized predicate is dominated by another verb.

It is said that Thai has two different so–called passive constructions: adversative passive and non-adversative passive whose details will be discussed in section 4. Examples of the two passives are provided below:

(5)   ADVERSATIVE:

    a. tamruat cap   khaw   ACTIVE
       police   catch   3p
       'The police arrested him.'

    b. khaw thuuk\doon (tamruat) cap   PASSIVE
       3Sg PASS police catch
       'He was arrested (by the police).'

(6)   NON-ADVERSATIVE :

    a. daawin chay [?upakoon baang yaang   ACTIVE
        Darwin use insturment some CL
        'Darwin used some instruments.'

b. ?upakoon baang [yaang chay dooy daawin]   PASSIVE
    instrument some CL use by Darwin
    'Some instruments were used by Darwin.'

We can see that the Thai non-adversative passive is more typical than the adversative passive since it mainly follows the definition of passive as presented in (4) in that, besides alternating with the active clause, its active object is promoted to be a subject, its active subject is demoted to be oblique as marked by the oblique marker *dooy*. However, for the analysis of the *thuuk/doon* passive as presented in (5), problems arise in that the string of verbs and the status of the Agent might lead us to consider it some construction other than a passive.

Unlike some languages such as English and French, where passives are obvious, the Thai "adversative passive" (sentence [5b]) derives via a process where the active object *khaw* is promoted and then occupies the subject position. The active subject *tamruat*, still preceding the active predicate, becomes optional. And the active predicate is preceded by *thuuk/doon,* marking for a passive reading. This fact results in the superficial similarity to the serial-verb construction as has been previously shown (sentence [1], [2], [3]). Therefore, one might argue that the adversative passive in Thai can be analyzed as a special case of serial – verb constructions and thus is biclausal rather than monoclausal since another verb *thuuk/doon* is added, making a string of verbs in the passive. Another problematic fact is that the Agent, if it appears, is in the usual subject position (preceding the active predicate) and is not overtly marked as oblique. The construction therefore should not be a typical passive.

Keenan (1985) also makes a claim for this characteristic of passive in Southeast Asia:

> Passives of this sort are widely attested in languages spoken in Southeast Asia, including Mandarin, although their analysis as passives is in fact not obvious. The languages which exhibit them independently are verb-serializing languages: apparently simplex sentences are commonly constructed with multiple verbs and few if any prepositions.

However, we will argue that the *thuuk/doon* passive construction has quite different syntactic and semantic properties from the typical serial verb construction in Thai. We will prove that the *thuuk/doon* construction should be considered a monoclausal and passive construction which reassigns grammatical relations within the clause, rather than a multiclausal serial- verb construction.

3. AN ARGUMENT FOR THE THUUK/DOON PASSIVE CONSTRUCTION.

Let us start with some examples of serial verb constructions and the *thuuk/doon* construction that are structurally similar:

(7)   SERIAL-VERB CONSTRUCTION :

        a. chan phuut set
           1p speak finish

        b. chan ?aw miit han nua
           1p take knife cut meat
           'I took a knife to cut meat.'

(8) thuuk/doon CONSTRUCTION:

        a. chan thuuk pron
           1p touch rob
           'I was robbed.'

        b. chan thuuk coon pron baan
           1p touch robber rob house
           'My house was robbed by a robber.'

The resemblance of the examples can be schematized as in (9) below :

(9) NP_1 V_1 (NP_2) V_2 (NP_3)

However, as we will discuss below, the examples in (7) and (8) have quite different syntactic and semantic properties, justifying their different structure classifications. We will present some evidence to show that the structure for the examples in (7) can be described as adjunct clauses, as in (10), and the examples in (8) are monoclausal and bear an embedded complement clause as the object of a matrix verb, as in (11). And it is the latter which is a passive reassignment of grammatical relations within the clause.

(10)   SERIAL VERB CONSTURCTION:

        NP_1 [V_1 (NP_2)] [V_2 (NP_3)]

(11)   *thuuk/doon* CONSTURCTION:

        NP_1 V_1 [(NP_2) V_2 (NP_3)]

In (10) the verb phrase headed by the first verb is an adjunct to the verb phrase headed by the second verb. By contrast, in (11) the first verb takes a clause or verb phrase as a complement, thus being considered monoclausal. Therefore, NP_2, if it occurs, in (10) has its grammatical role as an object of V_1, while NP_2 in (11) has its grammatical role as the subject of V_2. Moreover, V_1 in (11) is restricted to only *thuuk/doon*, while that of (10) has a much wider range. Additional evidence comes from extractibility, corresponding active clauses, deduction of NP_2 (active subject) and movement of NP_2.

3.1. Extractibility. Evidence that supports our argument that the *thuuk/doon* structure is monoclausal while other serial-verb constructions are multiclausal mainly comes from facts about extraction by negative placement and conjunction insertion.

In schematic (11), the negative marker *may* may appear before the first verb, but never between the two verbs as exemplified in (12) . By contrast, in schematic (10), the negative marker may occur either before the first verb or between the two verbs, as exemplified in (13) below:

      (12) a. chan may thuuk pron
          1p NEG touch rob
          'I wasn't robbed.'

        b. * chan thuuk may pron
          1p touch NEG rob

      (13) a. chan may phuut cop
          1p NEG speak end
          'I didn't finish speaking. (I didn't want to finish it.)'

        b. chan phuut may cop
          1p speak NEG end
          'I didn't finish speaking. (Something interrupted me.)'

We can see that the string of verbs in the *thuuk/doon* construction cannot be extracted by inserting the negative marker between the two verbs while that of the other serial-verb constructions can. This leads us to the conclusion that the string of verbs led by *thuuk/doon* is treated as a union rather than two separate verbs. On the other hand, following Law (1996), in a structure like schematic (11) the second verb, which is a complement of the first verb, is properly governed by the first verb, and the extraction is a violation. This is in contrast with the structure in (10) where the first verb is an adjunct to the second verb; neither verb governs the other. Therefore, the extraction in (13b) is not a violation.

Similarly, a string of verbs in the *thuuk/doon* construction cannot also be extracted by inserting a clausal conjunction (especially *con* 'until') while any other serial verb can, as exemplified below:

      (14) a. chan thuuk pron
          1p touch rob
          'I was robbed.'

        b. *chan thuuk con pron
          1p touch until rob

      (15) a. chan phuut cop
          1p speak end
          'I finished speaking.'

        b. chan phuut con cop
          1p speak untill end
          'I spoke until (I) finished.'

3.2. Corresponding Active Clauses. Another piece of evidence justifying the macro structures in (10) and (11) for the different types of constructions come from the fact that the multiclausal (10) does not have a corresponding active clause, from which the monoclausal (11) derives. The contrast is shown below:

         1   2   3   4
(16) a. chan thuuk coon pron
       1p touch robber rob
       'I was robbed by a robber.'

          3   4   1
   b. coon pron chan
      robber rob 1p
      'A robber robbed me.'

        1   2   3   4
(17) a. chan ?aw miit han
       1p take knife cut
       'I took a knife to cut (something).'

         3   4   1
   b. miit han chan
      knife cut 1p
      'A knife cut me.'

Example (16a) can be structurally changed (activized) to (16b) and still retain the same meaning. By contrast, suppose the example (17a) is structurally changed in the same way as (16a) is, the changed structure in (17b) does not semantically match (17a). Therefore, the possibility of deriving (17b) from (17a) is rejected. The clause must be considered a syntactically independent, not derived, construction. The reason supporting this is that (17a) consists of two separate clauses for which movement is impossible. Therefore, we can claim that the derivation of (16b) from (16a) is acceptable due to the fact that it is monoclausal.

3.3. Deletion of NP2. The next piece of evidence proving that *thuuk/doon* is not a serial-verb construction is that the NP_2 can be absent for *thuuk/doon,* but the NP_2 in serial verbs, occurring as the object of the preceding verb, cannot. Thus consider:

(18) a. chan   thuuk $coon_2$ pron
       1p   touch $robber_2$ rob
        'I was robbed by a robber.'

   b. chan   thuuk   pron
       1p   touch   rob
        'I was robbed.'

(19) a. chan   ?aw miit$_2$ han
        1p   take knife$_2$ cut
        'I took a knife to cut (something).'

    b.* chan ?aw han
        1p take cut

In the example (18b), NP$_2$ (*coon*) can be omitted since it is treated like the agent of a derived passive, while the NP$_2$ (*miit*) in (19b) cannot for it is the object of the transitive V$_1$.

3.4. Movement of NP$_2$. For *thuuk/doon*, NP$_2$ cannot be relativized and topicalized, and consequently fronted, since it is treated like a demoted active subject thus losing its grammatical relation. On the other hand, the NP$_2$ of serial verbs can. Thus consider:

(20) a. *coon thii chan thuuk pron
        robber$_2$ who 1p touch rob
        'the robber by whom I was robbed'

    b. *coon$_2$ chan thuuk pron
        robber$_2$ 1p touch rob
        'By the robber, I was robbed'

(21) a. miit$_2$ thii chan ?aw han
        knife$_2$ which 1p take cut
        'the knife which I took to cut meat'

    b. miit$_2$ chan ?aw han nua
        knife$_2$ 1p take cut meat
        'A knife, I took to cut meat.'

In (20), NP$_2$ (*coon*) cannot be relativized or topicalized because it is a demoted subject, whileNP$_2$ in (21) can since it functions like a typical object that can be both relativized and topicalized.

As we can see from the arguments above, the evidence shows that the *thuuk/doon* construction should be considered as monoclausal rather than adjunct clauses like other serial verb constructions. In this construction, the verb *thuuk/doon,* similar to *bei* in certain Chinese passive sentences (Dingxu Shi 1997), takes a special clause as its complement.[1] The subject of the complement clause is the agent NP and the predicate is what used to be called the passive verb. According to this analysis, the Thai

---

[1] This characteristic of passive also resembles that of Chinese, that is, the verb *bei* in certain Chinese passive sentences also takes a complement clause. Shi (1997) posits two *beis*. One *bei* is a preposition that takes the agent as its oblique object and the other is a passive marker that occurs right before the main verb:

    laoshi   bei   jiaole   toufa
    teacher   BEI   cut   hair
    'The teacher's hair was cut.'

adversative passive involves a complex sentence and is different from the passive in English or French. The absence of the agent NP in certain passive sentences is treated as a special case of null subject. As for the "oblique" status of the Agent or active subject, which might be opened to questioning since it is not grammatically marked as a typical oblique subject, we will argue that the active subject, in fact, has a syntactically peripheral "oblique" status. This claim is based on the observation of the grammatical limitations of the active subject: the deduction of the subject and the movement of the subject as previously shown. Therefore, even though not typical, this structure can reasonably be considered a passive on the aforementioned grounds.

4. Types of Passives.

As has been said by keenan (1985), if a language has two or more distinct passives, they are likely to differ semantically with respect to the degree of subject affectedness. This is also true in Thai. In Thai, passives vary according to whether the subject is positively or negatively affected. The negatively affected subject results in an adversative passive. While the positively affected subject results in a non-adversative passive, in Thai the adversative passive always carries the passive markers *thuuk/doon* while, in general, the non-adversative passive does not.

4.1. Adversative Passive or thuuk/doon Passive. As has been previously discussed in section 3, the adversative passive exhibits three syntactic properties: agent demotion, non-agent promotion and passive-voice marker. Since the basic active-transitive order in Thai is SVO, which can be extended as subject + verb + object + (oblique NP), to convert an active sentence into a corresponding adversative sentence, the patient object of an active is promoted and then fronted. Then the agent/dative subject of an active, remaining in the same position (before the verb of an active), is optional and limited in its grammatical functions. Finally, the active verb is dominated and marked for passive voice by *thuuk/doon,* as presented below:

> (22) a. rot chon dek   ACTIVE
>   car hit child
>   'A car hit a child.'

>   b. dek thuuk\doon rot chon   PASSIVE
>   child PASS car hit
>   'A child was hit by a car.'

> (23) a. khruu thamthoot nakrian   ACTIVE
>   teacher punish pupil
>   'A teacher punished a pupil.'

>   b. nakrian thuuk\doon (khruu) thamthoot   PASSIVE
>   pupil PASS teacher punish
>   'A pupil was punished (by a teacher).'

We can see that the derived verbs in the adversative passive construction are dominated by *thuuk/doon*. These verbs mean 'touch' literally. When either of them

functions as a passive marker, it gives the reading of 'undergo' (an unpleasant experience). The difference in the two verbs can be pragmatically defined as a matter of formality: *thuuk* is frequently found in formal contexts, while *doon* is less formally used, especially in less formal oral contexts.

4.2. Non-Adversative Passive. The passive of the non-adversative type is not common in academic textbooks. Prasithrathsint (1992) states that the non-adversative passive in Thai emerged as a result of Thai coming in contact with English. The passive of this type is quite different from the regular adversative passive in that 1) in general, the main verb occurs solely (The verb is not generally marked with any of the passive markers) and 2) the Agent is marked as a typical oblique, occurring in the proper position for an oblique. Even though the Patient of the non-adversative passive is also placed where the active subject normally occurs in an active sentence, the passive subject is interpreted not to be negatively affected at all. Consider the example of the non-adversative passive illustrated below:

(24) a. daawin chay [?upakoon baang yaang   ACTIVE
          Darwin use instrument some CL
          'Darwin used some instruments.'

     b. [?upakoon baang yaan] chay dooy daawin   PASSIVE
          instrument some CL use by Darwin
          'Some instruments were used by Darwin.'

However, it is not surprising that sometimes this passive structure is marked with *thuuk* when the subject is interpreted to be negatively affected, as shown:

(25) [?upakoon baang yaang] thuuk chay dooy daawin
          instrument some CL PASS use by Darwin
          'Some instruments were used by Darwin.'

In general, as has been previously discussed, the subject of the adversative passive is understood to be negatively affected rather than positively or neutrally effected whereas the subject of the non-adversative passive is understood to be positively or neutrally affected. It is odd to passivize certain sentences into adversative passives in which the subjects are supposed to be positively or neutrally affected. Thus consider the following clauses, where the negative affectedness on the subjects of clauses (26) – (28) is not evident, but for the clauses of (29)- (31) , no negative effect is apparent and the adversative is odd.

(26) [dek khon nan] thuuk\doon tii
          child CL that PASS hit
          'That child was hit.'

(27) maa thuuk\doon rot chon
          dog PASS car hit
          'A dog was knocked over by a car.'

(28) dek thuuk\doon thootthing
      dish PASS ignore
      'A child was ignored.'

(29) ? nakrian thuuk\doon chomchey
      pupil PASS admire
      'A pupil was admired.'

(30) ? noong thuuk\doon phii duulee
      younger PASS older take care
      'The younger sibling was taken care of by the older one.'

(30) ? [kawee khon nan] thuuk yokyong dooy [samaakhom nakkhian]
      poet CL that PASS admire by association writer
      'That poet is admired by a writer association.'

(31) ? [knwee khon nan] thuuk yokyong dooy [samaakhom nakkhian]
      poet CL that PASS admire by association writer
      'That poet is admired by a writer association.'

However, according to the prediction that *thuuk/doon* is beginning to be used with the neutral sense, many speakers use it with certain verbs without the adversative sense, as exemplified:

(32) khaw thuuk chuan pay ngaan
      3p PASS invite go party
      'He was invited to a party.'

5. POSSESSIVE RAISING IN THE ADVERSATIVE-PASSIVE CONSTRUCTION.

Another prominent feature of the Thai adversative passive is that, like the Chinese passive, the active object, when a possessive NP consisting of at least a possessed noun and a possessor, can be separated when the active sentence is converted into a corresponding passive sentence; namely, only the possessor is moved to occupy the subject position in the passive sentence, leaving the possessed noun in the former position. To discuss this issue, we need to understand Thai possessive constructions. An explanation will be presented in the following paragraphs.

In Thai the possessive construction is patterned as POSSESSIVD + (MARKER) + POSSESSOR. The possessive marker of the possessive construction is optional, especially in oral and/or informal contexts. When the possessive marker is absent, the possession can be indicated by its word order, for example:

(33) a. [pii khoong mee] suay cang
        sister POSS mother beautiful really
        'Mother's sister is really beautiful.'

    b. [pii mee] suay cang
       sister mother beautiful really
       'Mother's sister is really beautiful.'

    c. [mee phii] suay cang
       mother sister beautiful really
       'Sister's mother is really beautiful.'

In (33a), the possessor 'mother' is more obviously marked with the preceding possessive marker. In (33b,c), the inverted order shows different possessors and possessed head nouns.

When we passivize an active sentence where its patient object is a phrase containing a head noun and an adjacent possessor as its modifier, either the entire phrase or only the possessor can be promoted to be the subject of the corresponding adversative passive as exemplified in (34) below:

   (34) a. coon plon [baan khaw]   ACTIVE
       robber rob house 3p.
       'A robber robbed his house.'

    b. [baan khaw] thuuk\doon coon plon   NON-RAISING
       house 3p PASS robber rob
       'His house was robbed by a robber.'

    c. [khaw] thuuk \ doon coon plon [baan]   RAISING
       3p PASS robber rob house
       'His house was robbed by a robber.'

However, there is a restriction on promoting only the possessor. That is, only possessors that possess an inanimate thing can be promoted to the subject position, while those that possess an animate thing cannot.[2] Thus consider the examples below:

   (35) a. khaw te? [maa chan]   ACTIVE
       3p kick dog 1p
       'He kicked my dog.'

    b. [maa chan] thuuk \ doon khaw te?   NON-RAISING
       dog 1p PASS   3p kick
       'My dog was kicked by him.'

    c. * [chan] thuuk \ doon khaw te? [maa]   RAISING
       1p PASS 3p kick dog

---

[2] In other analysis (e.g. Shi 1997), the moved possessor is termed as the moved object and the retained possessed NP as the retained object.

6. SUBJECTHOOD IN THE PASSIVE CONSTRUCTIONS.

In general, subjects of non-basic sentences such as passives are less subject-like or do not have more subject properties than subjects of basic sentences since they are derived subjects which are converted from objects (Keenan, 1975). This is also true in Thai. The subject of a passive in Thai, besides taking the position of subject in the rigid SVO word order, can carry only some subject properties in behavior and control. Passive subjects can be controllers of coreference across clause boundaries, but to some extents. This limitation of the subjecthood in the passive constructions provides another piece of evidence that supports our argument.

6.1. Controllers of Coreference. Like the subject of an active sentence, the passive subject, in general, can control coreferential pronouns across clause boundaries. Thus consider:

(36) a. ACTIVE:

dam $_i$ toy deang $_j$ mua khaw $_i$ klap maa
Dam hit Deang when 3p return come
'Dam hit Deang when he came back.'

b. PASSIVE:

deang $_i$ thuuk dam $_j$ toy mua khaw $_i$ klap maa
Deang PASS Dam hit when 3p return come
'Deang was hit by Dam when he came back.'

The active subject *Deang* in (36a) generally controls the coreferential pronoun *khaw* in the following adverbial clause and so does the passive subject *Dam* in (36b). However, the control property of both subjects is not obvious enough, so the issue of pragmatics must also take part in the interpretation of the subject in such an adverbial clause.

6.2. Controllers of Coreferential Deletion. The subject of a passive can control coreferential deletion; that is, it can be coreferentially deleted across coordinate conjunctions. Thus consider:

(37) a. ACTIVE:

dam $_i$ toy deang $_j$ mua Ø $_i$ klap maa
Dam hit Deang when return come
'Dam hit Deang when coming back.'

b. PASSIVE:

deang $_i$ thuuk dam $_j$ toy mua Ø $_i$ klap maa
Deang PASS Dam hit when return come
'Deang was hit by Dam when he came back.'

Also, the deleted reference can refer backward:

(38)   a. ACTIVE :

mua Ø $_i$ klap maa dam $_j$ toy deang
when return come Dam hit Deang
'When coming back, Dam hit Deang'

b. PASSIVE:

mua Ø $_i$ klap maa deang $_j$ thuuk dam $_j$ toy
when return come Deang PASS Dam hit
'When coming back, Deang was hit by Dam.'

Unlike the controller of coreferential pronouns, the controller of coreferential deletion of both the subject of an active and subject of a passive is absolute. It is not likely for the coreferential deleted subjects in (37) and (38) above to be interpreted as other arguments.

However, there are some restrictions for the subject of a passive in this control property due to a semantic fact. The subject of passives cannot control the coreferential deletion if it is inanimate. The subject of a following adverbial clause in the chain, therefore, takes charge, being interpreted as an agent of the adverbial clause. Thus consider:

(39)   a. ACTIVE :

khonthay$_i$ phalit khaaw$_j$ phua Ø $_i$ song-?ook
Thai people produce rice for export
'Thai people produce rice for exporting.'

b. PASSIVE:

khaaw$_I$ phalit dooy khonthay$_j$ phua Ø $_I$ song-?ook
rice produce by Thai people for export
'Rice is produced by Thai people for exporting.'

7. PASSIVE VERSUS TOPICALIZATION.

The overt syntactic structures of passives in Thai generally look similar to topicalization in that both of their syntactically converted orders are the same (OSV) ; that is, the object of an active is topicalized and then fronted clause-initially. Thus consider the comparison between the active clause, topicalization clause, and passive clause below:

(40) a. khaw laa nuu duay kapdak   ACTIVE
        3p   hunt rat with   trap
        'He hunts rats with traps.'

b. nuu khaw laa ø duay kapdak TOPICALIZATION
  rat 3P hunt with trap
  'Rats, he hunts with traps.'

c. nuu thuuk/doon khaw laa duay kapdak PASSIVE
  rat PASS 3P hunt with trap
  'Rats are hunted by him with traps.'

Sentence (40b) is topicalization, with the Ø mark indicating the gap left in the sentence. Sentence (40c) is passive (of the adversative type) indicated by the passive marker *thuuk/doon*. Both types of sentences in Thai have the same sequence of derived constituents (PATIENT – AGENT – VERB). Thus, the only grammatical element that differentiates the two constructions is the passive maker. However, when we recall the non-adversative pattern without the passive marker, this pattern and topicalization are even more similar. Thus consider :

(41) a. khoong khaw hoo duay kradaat TOPICALIZATION
   thing 3p wrap with paper
   'That thing, (s)he wrapped with paper.'

  b. [?aakhaan nii] saang yaang dii NON-ADVERSATIVE
   building this build like good
   'This building was well built.'

Moreover, if a topicalization sentence is derived from an imperative clause, it represents the same pattern (PATTENT – VERB) as that of the non-adversative passive (42b), as shown below:

(42) a. hoo khoong duay kradaat IMPERATIVE ATIVE
   wrap thing with paper
   'Wrap the thing with paper.'

  b. khoong hoo duay kradaat TOPICALIZATION
   thing wrap with paper
   'That thing, wrap (it) with paper.'

Also, Prasithrathsint (1992) state that there is the absence of the adversative passive in some parts of Northeastern Thailand and shows that wherever a passive construction should occur, a topicalized construction would be used instead. This leads to the assumption that the Thai passive may arise through a reanalysis of topicalization or left-dislocation constructions. One could say that this agrees with the claim by Givon (1990b) that the left-dislocation construction overlaps functionally with the passive at the sub-domain of topicalized non-agent and with the passive syntactically in terms of word-order in SVO and SOV languages.

However, there are, according to Keenan (1985), some tests that can be used to distinguish the two similar structures synchronically. First, in passivization, the subject of an active clause is eliminated or relegated to the status of an oblique noun phrase in a way which the left-dislocations or the topicalizations, in general, do not. In the present paper, all the structures in which the subjects are eliminated thus will be

interpreted as passives. Therefore, sentences such as (43c), with the absence of an agent, should be considered passive rather than topicalized.

(43) a. ACTIVE:

boorisat yiipun phalit rotyon yaang dii
company Japan make automobile like good
'Japanese companies make automobiles well.

b. NON-ADVERSIVE:

rotyon phalit yaang dii dooy boorisat yiipun
automobile make like good by company Japan
'Automobiles are well made by Japanese companies.'

c. AGT ABSENCE:

rotyon phalit yaang dii
automobile make like good
'Automobiles are well made.'

(44) a. ACTIVE:

boorisat yiipun phalit rotyon yaang dii
company Japan make automobile like good
'Japanese companies make automobiles well.'

b. TOPICALIZATION:

rotyon boorisat yiipun phalit yaang dii
automobile company Japan make like good
'Automobiles, Japanese companies make (them) well.'

c. AGT ABSENCE:

? rotyon phalit yaang dii
automobile make like good

Next, major syntactic operations such as nominalization and relative- clause formation operate more freely on passive than on topicalizations. Thus consider the examples below where a passive sentence in Thai can be nominalized and relativized, as in (45), while a topicalized sentence cannot, as in (46):

(45) a. NON-ADVERSIVE:

rotyon phalit dooy boorisat yiipun
automobile make by company Japan
'Automobiles are made by Japanese companies.'

b. NOMINALIZATION:

kaanthii rotyon phalit dooy boorisat yiipun
NOM automobile make by company Japan
'That automobiles are made by Japanese companies.'

c. RELATIVIZATION:

muan thii rotyon phalit dooy boorisat yiipun
city which automobile make by company Japan
'the city where automobiles are made by Japanese
companies'

(46) a. TOPICALLZATION:

rotyon boorisat yiipun phalit
automobile company Japan make
'Automobiles, Japanese companies make.'

b. NOMINALIZATION:

* kaanti rotyon boorisat yiipun phalit
NOM automobile company Japan make

c. RELATIVIZATION:

*muang thii rotyon boorisat yiipun phalit
city which automobile company Japan make

Moreover, when passivizing, modal verbs no longer accompany derived verbs while they always attach to main verbs in topicalization. Thus consider:

(47) a. ACTIVE:

thea kamlang sangsoon luuk
3p PROG teach son/daughter
'She is teaching her child.'

b. PASSIVE:

*luuk thuuk\doon thea kamlang sangsoon
son/daughter PASS 3p PROG teach

(48) a. ACTIVE:

dek-dek kamlang kin khanom
children PROG eat dessert
'The children are having dessert.'

b. TOPICALIZATION:

khaom dek-dek kamlang kin
dessert children PROG eat
'Dessert, the children are having (it).'

8. THE TEXT DISTRIBUTION OF ACTIVES AND PASSIVES.

In Thai the functional distribution and overall frequency of Active and passive clauses in discourse is different. As stated by Givon (1990b), humans tend to topicalize or talk about the more agent-like participant, thus active clauses appear more frequently in text than the two passives do. To illustrate this argument, the frequency distribution of actives and passives in Thai is shown below:

TABLE 1. RELATIVE FREQUENCY OF ACTIVES AND PASSIVES IN THAI TEXT[3]

Construction type	More N	academic %	Less N	academic %	Overall N	%
Active						
Intransitive	277	55.4	320	64	597	59.7
Transitive	181	36	168	33.6	349	34.9
Passive						
Adversative	19	3.8	3	.6	22	2.2
Non-adversative	23	4.2	9	1.8	32	3.2
Total	500	100	500	100	1,000	100

As shown in the data, active clauses, predominate overall (59.7 % + 34.9 % = 94.6 %). Both the adversative and non-adversative passive seem to occur more frequently in academic texts, such as a textbook in anthropology by Terwiel, Diller, and Satayawatana (1990), (8 %), more than in less academic and formal texts, such as a casual-style sermon by Phuthathatphiku (n.d.), (2.4 %). The non-adversative passive appears more frequently, but only slightly, than the adversative in both more academic and less academic texts.

9. CONCLUSION.

Thai presents two types of passives, organized mainly according to the affectedness of the subjects of passives. The two passive types are labeled as the adversative passive and non-adversative passive. Structurally, the adversative type is not prototypical, like in English and French. Similar to Chinese, it is marked by the verb of reception meaning 'touch' literally and thus its subject is negatively effected.

---

[3]Terwiel, Diller and Satayawatana (1990) was chosen for more academic texts and Phuthathatphiku (n.d.) for less academic texts since they show a significant difference in degrees of academic qualifications, that is, the former is a textbook in theoretical anthropology while the latter an informal monologue. In each text, the first 500 sentences were taken for our identification.

In general, the non-adversative type is lexically unmarked and not negatively effected, but it is more prototypical since it is believed to come from English and French. In general, the passives look structurally similar to other constructions such as serial verbs and topicalized clauses, but there are a number of pieces of evidence that can be used to distinguish the passives from other constructions in Thai.

REFERENCES

_____

KEENAN, EDWARD L. 1985. Passive in the World's Languages. *Language Typology and Syntactic Description Vol. 1*, ed. by Timothy Shopen, 243-281. Cambridge: Cambridge University Press.

_____. 1975. Some Universals of Passive in Relational Grammar. *Chicago Linguistic Society.*

GIVON, T. 1992. *Syntax : A Functional-Typological Introduction Vol.1.* Amsterdam: J.Benjamins.

LAW, PAUL. 1996. A Note on the Serial Verb Construction in Chinese. *Cahiers de Linguistique – Asie Orientale* 25(2).199-233.

LAKAYAMA, TOSHIHIDE. 1997. Functions of the Nootka (Nuu-Chah-Nulth) "Passive" Suffix. *International Journal of American Linguistics* 63(3).412-31.

PRASITHRATHSINT, AMARA. 1990. *Phasasatsankhom* (*Sociolinguistics*). Bangkok: Chulalongkorn University.

PHUTHATHATPHIKHU. (n.d.). *Silapa Heng Kan Mi Chiwit Yoo Duay Patchuban* (*The Art of Living in the Present*).

SHI, DINGXU. 1997. Issues on Chinese Passive. *Journal of Chinese Linguistics* 25(1). University of California at Berkeley.

SOOKGASEM, PRAPA. 1992. A Verb-Subject Construction in Thai: An Analysis of the Existential Verb *mii*". *Pan-Asiatic Linguistics Vol. 1*. Bangkok: Chulalongkorn University.

TERWIEL.B.J., DILLER, ANTHONY and SATAYAWATANA, CHONTIRA. 1990. *Khon Tai (Dom) May Day Yoo Thi Ni* (*Tai People Did Not Live Here Before.*) Bangkok : Meang Boran.

THE ASPECTUAL SYSTEM OF CHIYAO[*]

ALFRED J. MATIKI

University of New Mexico

1. INTRODUCTION.

This paper explores the aspectual system of Chiyao, a Bantu language spoken in parts of Malawi, Mozambique, and Tanzania, and classified by Guthrie (1971) as belonging to the Zone P language group together with Mwera, Makonde, and Makua. The distinction between aspect and tense is not always easy to draw clearly in individual languages (cf. Langacker 1972:207; Comrie 1976:82-3; among others). As a matter of fact, some scholars have gone to the extent of making a distinction between aspectual languages and non-aspectual languages while others have distinguished between aspect and *Aktionsart* (German for, 'kinds of action'), "with aspect referring to instances where the opposition has been grammaticized, *Aktionsart* to instances where it has been lexicalized" (Crystal 1997:29-30). A more radical view was expressed by Zandvoort (1962) who, using the classical examples of the aspectual system of Russian, claimed that aspect is an exclusive phenomenon for Slavonic languages. This contention appears to be based on the notion that aspect can only be expressed morphologically. Zandvoort's analysis, therefore, excludes languages that do not inflect for or morphologize aspect. It has, however, become apparent in studies subsequent to Zandvoort's claims that the so-called non-aspectual languages can and do express aspectual concepts by other linguistic means (cf. Hoepelman 1974, Comrie 1976, Melchert 1980, Bybee 1985).

This paper will therefore assume the position stated above, namely, that aspect can also be expressed by other linguistic means. In this connection, it is important to adopt a definition of aspect which is not specific to the peculiar morphosyntactic configurations of particular languages. A definition which fits this criterion is the one elaborated by Comrie (1976:3), namely, that "aspects are different ways of viewing the internal temporal constituency of a situation". For instance, a situation could be viewed from the point of view of its beginning (inception), its middle phase (continuation), or its completion, among other points of view. In a similar vain, Langacker (1972: 207) points out that "aspect refers to the inception, duration, or completion of an event". This definition, however, is limited in that it does not allow for such aspects as iterative and habitual. This paper will, therefore, attempt to describe some of the aspectual distinctions that obtain in Chiyao and to relate these distinctions to aspect as a general linguistic category. The data used in this paper are drawn from my own intuitions as a native speaker of Chiyao.

2. GENERAL REMARKS ABOUT CHIYAO.

Structurally, Chiyao resembles other Bantu languages with agglutinative word morphologies. Every part of speech consists of a stem or root and various prefixes and suffixes which modify the primary meaning of the root. Nouns belong to various nominal classes which determine agreement patterns when in construction with other phrases. The verb is morphologically complex and consists of a verbal radical and a number of affixes and suffixes which encode agreement markers, tenses, negation, causatives, applicatives, statives, reciprocals, and other verbal extensions.

Chiyao, like many other Bantu languages, is a tone language. While nouns are underlyingly specified for tone, verbal tone is, for the most part, dependent on the tense-aspect morphology of the verb. As Mtenje (1993:180) notes, aspectual and tense markers trigger the assignment of high tones to certain domains in the verb through a number of independently motivated rules. In this paper, however, tonal and phonological phenomena in general are not discussed, as they are not particularly pertinent to the current discussion. Those interested in these phonological issues of Chiyao should see Mtenje (1990a, 1990b, and 1993).

3. TENSE DISTINCTIONS IN CHIYAO.

As in many other languages, there is evidence in Chiyao of the three devices that usually signal aspect, namely, inflectional or derivational morphology, lexical items, and sentential constructions. All these devices attest in Chiyao with different degrees of productivity and frequency. Before we examine these aspectual distinctions, however, it is important to briefly review the tense distinctions that obtain in Chiyao. This is important because there are a lot of cases where it is difficult to distinguish between tense and aspect since they both deal with time. Tense is a "deictic category that places a situation in time with respect to the moment of speech, or occasionally with respect to some other pre-established point in time" (Bybee 1985:21) while aspect relates the time of an event to the internal constituents of the event. The distinction between tense and aspect is also referred to in some literature as situation-external-time and situation-internal-time, with the former referring to tense and the latter to aspect (see Comrie 1976:5).

Chiyao has a five tense system involving three major contrasts –present, past, and future. The past tense divides into simple past and remote past while the future also divides into simple or immediate future and remote future. These contrasts are briefly elaborated on in the sections that follow.

3.1. Present Tense (PRES). The present tense describes situations which are presently taking place or which will be taking place in the immediate future. This tense form is formed in Chiyao by prefixing the agreement marker to the infinitive form of the verb as the following two examples show.

(1)   Nguku   si-ku-sa-lya   mpunga.
        chicken they-PRES-HAB-eat   rice
        'Chickens eat rice.'

(2)   Tu-ku-tawa   nyumba.
        we-PRES-build house
        'We are building a house.'

Although we will defer the discussion of the aspectual content of these examples, it is clear that the present tense in Chiyao, like in a lot of other languages, carries various notions of aspect with the moment of speech as the reference point.

3.2. Past Tense (PAST). The past tense indicates a situation which occurred before the moment of speech. The past tense, too, combines with various aspects. As we noted above, Chiyao distinguishes between simple or near past and remote past. The simple past is the unmarked member in this tense group and refers to situations which existed or took place before the present moment, either earlier today or yesterday. It is formed by prefixing an agreement marker to a perfective verb stem. Example (3) illustrates the simple past, while example (4) illustrates the scope of the simple past, namely that it cannot be used for events that took place before yesterday.

(3)   Tu-piikeene   kuti   a-yiche   lelo/liso.
        we-hear.PERF   that   he-come.PERF today/yesterday.
        'We (have) heard that he came today/yesterday.'

(4)   *Tu-piikeene   kuti   a-yiche   lijusi.
        we-hear.PERF   that   he-come.PERF the day before yesterday
        'We (have) heard that he came the day before yesterday.'

The remote past is used for situations which existed or took place earlier than yesterday. It is formed by prefixing the remote formative *a-* to the perfective stem after the agreement prefix as illustrated in example (5). Example (6), on the other hand shows that the remote past cannot be used for events that took place today although in some dialects it may be used for events that occurred yesterday together with the simple past.

(5)   Tu-piikeene   kuti   wa-a-yiche   lijusi.
        we-hear.PERF   that   he-REM-come.PERF   the day before yesterday
        'We heard that he came the day before yesterday.'

(6)   *Tu-piikeene   kuti   wa-a-yiche   lero.
        we-hear.PERF   that   he-REM-come.PERF   today
        'We heard that he came today.'

3.3. Future Tense (FUT). Chiyao divides the future tense into the immediate future and the remote future. The former is realized by prefixing the formatives *ti- ,ta-*, or *tu-*, depending on the number and nature of the subject NP, to the agreement prefix and verb

stem. The immediate future refers to situations which will occur subsequent to the present moment but during the same day, the same week, the same year or any other immediate single block of time as example (7) shows.

(7)   Ti-n-dawe   nyumbalelo/namyaka.
        FUT-I-build   house   today/this year
        'I will build a house today/this year.'

The remote future refers to situations which will occur after today or after this year and is marked by a remote formative *ji-* for the singular and *chi-* for the plural. The following example illustrates this future tense.

(8)   Ti-n-ji-tawa   nyumbachaka cha mkucha.
        FUT-I-REM-build   house   the year after next year
        'I will build a house the year after next year'

Like the present tense, the future tense has other functions beyond indicating futurity. It also indicates intention, obligation, and a confident prediction that the event will be actualized. The future marker in Chiyao and many other languages is also used as an authoritative alternative to the imperative form. These issues, however, are beyond the purview of the current discussion.

4.   ASPECTUAL DISTINCTIONS IN CHIYAO.

We will now turn to look at the various aspectual distinctions that obtain in Chiyao. As in many other languages, Chiyao uses three devices to signal aspect - inflectional or derivational morphology, lexical items, and sentential constructions. Both perfective and imperfective aspects are evident in Chiyao.

4.1. Aspects Related to the Perfective (PERF). Perfective aspect 'involves lack of explicit reference to the internal temporal constituency of a situation' (Comrie 1976:21). Chiyao marks perfectivity by suffixing a perfective formative to the verb stem and changing the final vowel (FV) from *-a* to *-e*. The actual shape of the formative varies depending on its phonological environment. Some of these variant formatives are *-il-, -it-,* and *-ich-* and are illustrated in the following derivations and sentential examples.

(9)   a.   -lola 'see' > -lol-it-e 'have seen'
        b.   n-dol-it-e
                I-see-PERF-FV
                'I have seen.'

(10)   a.   -tawa 'build' > -taw-il-e 'have built'

           b.   Maswani   a-taw-il-e   nyumba   jakusalala.

                     Maswani   he-build-PERF-FV   house   it.beautiful

                     'Maswani has built a beautiful house.'

(11)   a.   -utuka 'run' > -utw-ich-e 'have run'

           b.   Juma   wa-a-utw-ich-e   mjaha liso

                     Juma   he-REM-run-PERF-FV race   yesterday

                     'Juma (had) ran a race yesterday.'

(12)   a.   -tanda 'start' > -tand-it-e 'have started'

           b.   Juma   wa-a-tand-it-e   kujiganya.

                     Juma   he-REM-start-PERF-FV   teaching

                     'Juma (had) started teaching.'

It will be noted that the forms presented above are quite similar to the past tense forms illustrated earlier because, as it was stated then, the past tense form is derived from the perfective. This is easy to understand since perfective aspect denotes "a complete situation, with beginning, middle, and end" (Comrie 1976: 18). It is for this reason that perfective aspect is, for the most part, restricted to past situations. Furthermore, the past tense is intrinsically perfective.[1] When the perfective is used with the future tense ('future in the past') it uses an auxiliary construction, probably an indication that perfectivity and futurity are not natural partners. The following sentence exemplifies perfective aspect used with the remote future tense. We ignore the final vowel (FV) in this example and in all subsequent ones, as it is not pertinent to the discussion.

(13)   Juma   ta-a-chi-wa   a-li   a-utwiche

           Juma   FUT-he-REM-be   he-be he-run.PERF

           'Juma will have ran (by that time).'

The examples of the perfective aspect shown above (10 - 12) also illustrate a distinction between completive and achievement aspect. The distinction between these two is lexically expressed. Completive notions of aspect are expressed by telic predicates; these are predicates that describe situations that have an intrinsic time continuum and have a built-in end point. The examples in (10) and (11) above demonstrate this completive notion. Even though the act of building a house or running a race can be seen as durative, in the sense that it has a starting point and an end point, the focus of the

---

[1] What is strange about perfective aspect in Chiyao, however, is that the perfective formative is realized as a suffix rather than as a prefix as is the case with tense markers and other aspectual markers such as habitual aspect. This trend is also evident in other languages closely related to Chiyao such as Mwera. Other Bantu languages, such as Chichewa, have a different pattern suggesting that this positioning of the perfective formative is an idiosyncratic phenomenon of Bantu languages in group 20 of Zone P. Joan Bybee (personal communication) suggests that the positioning of the perfective probably has to do with the fact that this formative is reconstructed for Proto-Bantu as -ile-, a verb meaning 'finish', and is iconically suffixed to the verbal radical; one does the action first and finishes it. It is very likely that the perfective formative was originally derived from an adverb (see Bybee et al. 1994).

predicate here is on the expression of an event that proceeded to an end point. Thus, the sentences in (10) and (11) entail the existence of complete products, a house and a race, respectively, rather than the processes involved in accomplishing these end products.

Achievement predicates differ from accomplishment predicates in the sense that the former are non-durative by nature. These, too, are intrinsically perfective and include such verbs as *-pagwa* 'be born'*, -simana* 'find'*, -tanda* 'begin/start' and so on. Example (12) above also illustrates this form.

4.2. Aspects Related to the Imperfective. The imperfective is the opposite polar of the perfective and is characterized by Comrie (1976:24) as having an "explicit reference to the internal temporal structure of a situation, viewing a situation from within." Unlike the perfective which views a situation as a bounded single entity with a beginning and an end, imperfective aspects view a situation as on-going and therefore unbounded. It is for this reason that in some literature perfective aspect is also called completive aspect while imperfective aspect is called non-completive (see Machobane 1985). This nomenclature is, however, a little confusing and is not used in this paper in that sense. Aspects related to the imperfective found in Chiyao include the progressive, habitual, and iterative.

4.2.1. Progressive Aspect (PROG). Progressive aspect indicates continuous action and is marked by the prefix *ku-* attached to the verb stem in the present tense. In the past and future tenses, the progressive formative is suffixed to the stem after the final vowel and is realized as *-je* for the simple past and simple future and *-ga* for the remote past and remote future.[2]

It will be noted below that some of the examples of the progressive aspect in the past and future tenses have been contextualized with an adverbial; otherwise the forms would have a strong habitual-progressive interpretation rather than a progressive interpretation only. In the context of the present study, however, we will regard these formatives as progressive mainly because that is the main interpretation that they give to the predicate stem to which they are attached. The following examples illustrate this aspect in all the tenses.

> (14) Juma a-ku-papula   buuku.
>   Juma he-PROG-shred   book
>   'Juma is shredding a book.'

---

[2] The real status of the progressive formatives for the past and future tenses needs further investigation. Their position in the verbal morphology is a little strange for suffixes. Suffixes, such as valence changing suffixes, in Chiyao and probably other Bantu languages, usually come between the stem and the final vowel. It is usually abbreviated adverbials that attach to the verb after the final vowel as illustrated in this example:

> Juma   ali   m-ku-utuka-pe.
> Juma   he.be   in-PRES-run-still
> 'Juma is still running.'

The past and future progressive formatives are probably recent formations and likely to have been derived from adverbials.

(15)　Juma　a-utwiche-je.
　　　　Juma　he-run.PERF-PROG
　　　　'Juma was/has been running.'

(16)　Juma　wagwile　　　pa-wa-a-utuka-ga.
　　　　Juma　he.fell.PERF　when-he-REM-run-PROG
　　　　'Juma fell when he was running.'

(17)　Juma　ta-a-utuche-je　　　pankawapa.
　　　　Juma　FUT-he-run-PROG　　soon.very
　　　　'Juma will be/start running very soon.'

(18)　Juma　ta-a-chi-utuka-ga　　　pandawijo.
　　　　Juma　FUT-he-REM-run-PROG　　at.time.then
　　　　'Juma will be running at that time.'

Although we have restricted our examples to activity verbs, it should be noted that stative verbs in Chiyao also take the progressive marker (examples 19-20). Unlike activity verbs, however, stative verbs with a progressive marker do not receive a progressive interpretation. The progressive marker in such constructions simply represents the unchanging nature of the states; "unless something happens to change the state, the state will continue." (Comrie 1976:49).

(19)　Juma　a-ku-lwala.
　　　　Juma　he-PROG-ill
　　　　'Juma is ill.'

(20)　Juma　a-lwasile-je.
　　　　Juma　he-ill.PERF-PROG
　　　　'Juma was/has been ill (he is fine now).'

It should be noted here that the progressive *ku-* in isolation is usually interpreted as the infinitive marker, probably derived from a locative preposition. However, in the context used above (14), it is unquestionably progressive in its interpretation. Another important distinction that has to be made here is whether this formative *ku-* is primarily a tense marker and, in which case, the progressive aspect is unmarked in this tense or whether it marks both tense and the progressive aspect. We will be able to provide an answer after examining our next aspectual notion.

4.2.2. Habitual Aspect (HAB). Habitual aspect describes 'a situation which is characteristic of an extended period of time, so extended in fact that the situation referred to is viewed not as an incidental property of the moment but, precisely, as a characteristic feature of a whole period' (Comrie 1976: 27-28). Habitual aspect in Chiyao is marked by

sa- in present tense verbs. The habitual in the other tenses has already been alluded to in the examples in (15) through (18) when used without the adverbials. As stated earlier, however, the habitual aspect in these two tenses also takes a progressive interpretation. Examples (21) through (23) show the habitual aspect. Example (24), on the other hand, shows that the habitual marker may sometimes be used to express a universal truth rather than habituality. This is not strange to the natural languages of the world as the habitual aspect is often used for the expression of timeless truths (Lyons 1977: 716, Bybee et. al. 1994).

(21)   Juma   a-ku-sa-utuka   nlope.
        Juma   he-PRES-HAB-run   very
        'Juma runs very fast.'

(22)   Juma   wa-a-mwa-ga   ukana.
        Juma   he-REM-drink-HAB   beer
        'Juma used to drink beer.'

(23)   Juma ta-a-chi-mwa-ga   ukana   pataachikula.
        Juma FUT-he-REM-drink-HAB   beer   when.FUT.he.REM.grows
        'Juma will be drinking beer when he grows up.'

(24)   Achim'masyeeto   a-ku-sa-weleka   wanache.
        women   they-PRES-HAB-bear   children
        'Women give birth to children.'

One of the issues we raised above was whether *ku-* primarily signals tense or aspect in Chiyao. Although aspect is generally regarded as more basic than tense given that children seem to acquire the former more quickly than the latter (Lyons 1977:705), the case of *ku-* appears to be different. The presence of *ku-* in the present tense form in (21) and its absence from the past and future forms seems to confirm somewhat the speculation we raised above that this formative, although it has both a present tense interpretation and a progressive one as well, is primarily used to signal the present tense. This is not an unusual phenomenon as we have already noted that tenses convey various aspectual notions. It should also be noted that a lot of languages do not combine progressive aspect and habitual aspect. This explains why the *ku-* in (21) and (24) is glossed as present tense rather than progressive aspect.

4.2.3. Iterative Aspect. Iterative aspect describes 'the repetition of a situation, the successive occurrence of several instances of the given situation' (Comrie 1996: 27). Chiyao marks iteration iconically by reduplication of the verb stem. The choice of reduplication for iteration is not particularly surprising. As Bybee (1985: 150) points out, iteration is usually derivational and not fully integrated into the inflectional aspectual system. In Chiyao, though, all tenses accept iteration and it also combines very easily with the other aspectual forms. In (25) below we show examples of reduplicated forms.

In (26) we illustrate the fact that some verbs are inherently iterative and do not need to undergo any reduplication.

(25) Reduplicated Forms

-utuka	'run'	⇒	*-utukautuka* 'run many times/often',
-walanga	'read/count'	⇒	*-walangawalanga* 'read/count often',
-sulula	'drip/leak'	⇒	*-sulula-sulula* 'drip/leak often',
-lya	'eat'	⇒	*-lya-lya-lya* 'eat and eat and eat'
-jenda	'walk/travel'	⇒	*-jendajenda* 'wander about/travel often'

(26)   *-tetemela* 'tremble'   *-kupila* 'blink', *-sulula* 'drip/leak'.

Reduplicated forms may also combine with the habitual marker and have the attribute of willfulness. For instance, the example in (27) has the added meaning that the uncle makes a certain degree of commitment to being ill.

(27)   Mjomba wangu   a-ku-sa-lwala-lwala.
        uncle   he.my   he-PRES-HAB-ill-ill
        'My uncle gets ill often.' or 'My uncle makes a habit of being ill.'

5. SUMMARY.

We have demonstrated in this paper that tense and aspect are intricately connected in Chiyao and that the latter is signaled through inflectional, derivational, and lexical devices. Like aspect in other languages, the different aspectual distinctions that obtain in Chiyao cross-classify with tenses, except that it is most common for the perfective to be restricted to the past tense. As we noted, the past tense is formed from a perfective stem, making it intrinsically perfective. The progressive and iterative aspects easily combine with all tenses while the habitual aspect co-occurs with all tenses except the simple past and the simple future. We have also noted, for instance, how the suffixes *-ga* and *-je* are used to signal both progressive aspect and habitual aspect and how they are linked to the past and future tense. Another formative that appears to have several functions is *ku-*, signaling present tense, progressive aspect, and stativity, apart from being a locative preposition and an infinitive marker.

We have also noted that the position of aspectual markers with respect to the verbal stem is highly variable, unlike the other members of the verbal complex such as tense, valence, agreement markers and so on. There are aspectual markers that are prefixed to the verb stem, those that are suffixed to the stem before the final vowel and those that are suffixed after the final vowel. These facts do not seem to support the 'tense before aspect' hypothesis (Bybee 1985) and the generally fixed positions of affixes in verbal complexes.

Bybee (1985) shows that in many languages inflectional affixes occur in a predictable order away from the verbal radical. For instance, aspectual markers are usually closer to the verbal root than tense markers. We have already explained why the order of aspectual

affixes with respect to the verbal radical in Chiyao is variable, namely that they are probably a recent formation in the language and are still in the process of grammaticization. This is not particularly surprising given that languages that make a distinction between simple/near and remote tenses do not generally have aspectual markers (Aaron Smith, personal communication). The various positions of aspectual affixes in Chiyao also reflect their multiple sources, and we

have already alluded to both adverbials and prepositions as possible sources of these affixes.

REFERENCES

BYBEE, JOAN L. 1985. *Morphology: A Study of The Relation Between Meaning and Form*. Amsterdam: John Benjamins.

BYBEE, JOAN, REVERE PERKINS, and WILLIAM PAGLIUCA. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.

COMRIE, BERNARD. 1976. *Aspect*. Cambridge University Press

CRYSTAL, DAVID. 1997. *A Dictionary of Linguistics and Phonetics*. Oxford, UK: Blackwell Publishers.

GUTHRIE, MALCOLM. 1971. *Comparative Bantu 2*. Farnborough, England: Gregg International Publishers.

HOEPELMAN, J. 1974. Tense, Logic, and the Semantics of Russian Aspect. *Theoretical Linguistics* 1(1-2).158-180.

LANGACKER, RONALD W. 1972. *Fundamentals of Linguistic Analysis*. New York: Harcourt, Brace, Jovanovich.

LYONS, JOHN. 1977. *Semantics II*. Cambridge: Cambridge University Press.

MACHOBANE, M. 1985. *Tense and Aspect in Sesotho*. Indiana University Linguistics Club Monograph.

MELCHERT, H. C. 1980. Some Aspects of "Aspect" in Mandarin Chinese. *Linguistics* 18(7-8).635-654.

MTENJE, AL. 1990a. Feature Spreading in Phonology: Evidence from Chiyao. *Linguistic Analysis* 20(1-2).87-103.

MTENJE, AL. 1990b. Verbal Tone in Chiyao. Paper read at the 21st Annual Conference on African Linguistics, University of Georgia.

MTENJE, AL. 1993. Verb Structure and Tone in Chiyao. *Topics in African Linguistics*, ed. by Salikoko S. Mufwene and Lioba Moshi, 179-190. Amsterdam and Philadelphia: John Benjamins Publishing Company.

ZANDVOORT, P. W. 1962. Is "Aspect" an English Verbal Category? *Contributions to English Syntax and Philology,* ed. by Frank Behre. Gothenburg Studies in English, 14.

ADDRESS FORMS IN CHINESE AND THEIR INTERACTIONAL FUNCTIONS[*]

JUN YANG

The University of Arizona

1. INTRODUCTION.

The purpose of the present paper is two fold: (1) to describe address forms in Chinese, and by so doing, highlight the characteristics of Chinese address forms; and (2) to highlight the functions these address forms perform in interpersonal interaction. It will be shown that in Chinese, kinship terms are used extensively as address forms for both close and distant relationships, and perform a special social function in the ritual of inducting the young into the culture. It will also be shown that address forms in Chinese are constantly chosen with care by speakers as strategies to achieve personal ends in interactions.

2. UNIVERSALITY OF ADDRESS FORMS ACROSS LANGUAGES.

Brown (1965), after investigating 27 languages in the world, came to the conclusion that social events like addressing have great regularity, that is, there is a universal relationship between address forms used and the status and solidarity they indicate, a conclusion known as the "invariant norm of address". The norm stipulates that "the forms of address used by the superior of an unequal dyad are also the forms reciprocally used between intimate equals, whereas the forms used by the inferior of an unequal dyad are also the forms reciprocally used between distant equals" (Kroger and Wood 1992, 149).

Chinese was among the languages investigated by Brown and shown to support the claim of the "invariant norm of address", but I believe the supporting result was an artifact of modern Chinese which uses personal pronouns as address terms as well. As I see it, the observation of universality holds true for European languages that mainly use pronouns, e.g. Vous and Tu in French or personal names such as the full name or title plus last name used in English. In a dyad encounter, the forms of address chosen in Chinese may well index status and solidarity. However, the Chinese address system, because of its extensive use of kinship terms which are in nature not reciprocal, does not support this generalization. Another reason that Brown's conclusion is dubious is that solidarity and status may be different concepts in different cultures. For example, in the US, politeness and friendliness correlate highly and could be treated as a single index, but in Japan they are distinct. (Ide, S., Ikuta, S., Kawasaki, A., and Ogina, T. 1992). Japanese speakers' choice of politeness strategies is to a large extent decided by the consideration of power relations while Americans adopt strategies consonant with solidarity. (Ellis

---

[*] Address form and address term are used interchangeably in the present paper.

1994, 180). In this paper, I will describe the use of Chinese address terms with regard to the concept of politeness in Chinese and explore the multiple functions of addressing.

Although Brown's conclusion does not hold for all languages, the three indices proposed for dyad interaction are still useful tools of analysis. It is now almost a given that any given society will have some equal and non-equal, solitary and non-solitary dyads, and each language has its own means of managing differences and changes in status and solidarity as reflected in the address terms chosen. In this paper, Chinese address terms will be analyzed in terms of the three indexes proposed by Brown. They are the reciprocity of the exchange, the solidarity of exchange and the direction of unequal exchange.

3.   CONFUCIAN THOUGHT AND THE USE OF ADDRESS FORMS.

An account of the use of address forms in Chinese cannot be complete without an account of the concept of politeness in Chinese culture. In modern Chinese, the word for politeness is "li mao". "Li" is a key concept of Confucianism, which means "hierarchy". That a word implying "social hierarchy" comes to be used to mean politeness is not by accident (Gu 1990, 237). In orthodox Confucian thought, naming practice has a very important role to play in regulating a society. According to the teaching of Confucius, for the harmony of the universe to be achieved and maintained, of which human society is an integral part, it is necessary for everybody in the society to have a proper place. This assignment of position in the society is achieved by giving everyone a proper name, a practice known as "zheng ming", which means "correct names". A name is not just a label, for it signifies a relationship. "Zheng ming" is to put each individual in his/her place according to his/her social position, and speech has to be used in accordance with a user's status in the social hierarchy. The idiom "ming bu zheng, ze yan bu shun" means that if one does not have the names correct, he/she cannot speak properly, not mention to behave appropriately. "Zheng ming" was taken so literally and followed so dogmatically that taboos about names evolved. For example, any character in the given name of an emperor was not to be used by anyone else. Even in their daily life, people had to take extreme care not use that character, not even the character that has the same or similar pronunciation. The consequence could be imprisonment or death penalty. An emperor's name remained a taboo even after he died. In traditional Chinese culture, the use of names was closely related with social hierarchy and the Chinese idea of politeness.

Interpersonal relationships in China can be divided into two types: (1) intimate and close, and (2) distant and outside. As Scollon and Scollon (1991, 118) observed, "the Asian world is divided into inside (nei) and outside (wai) relations. Inside relations are those of close regular contact, such as family members, friends, co-workers and schoolmates. Outside relations are those temporary contacts one has with taxi-drivers or clerks in service encounters." This is certainly true of China. In dealing with outsiders or strangers, Chinese tend to be extremely polite, so polite that sometimes it is confusing to a foreigner. For example, an exchange student from the US, in perfect shape, sometimes found himself being offered a seat on the bus. In addition, Chinese speakers "tend to

focus attention on their relative position vis-à-vis the addressee's" (Pan 1995a, 249) when choosing the right interactional strategy. This is not only true of interaction with outsiders but also true with insiders. To teach people better manners, Confucius particularly emphasized five relations, known as "wu lun" which means "the five orders". They are relations exemplified between ruler/subject, father/son, elder brother/younger brother, husband/wife, and friend/friend (Young 1994, 151). These five orders with the exception of the last one are by definition hierarchical. This partly explains why even among relations of solidarity, Chinese will often make a power distinction, which seems to be paradoxical. The first order relationship, the most important order of relationship, dictates that the ruling power or the institutional power associated with it has the overriding power in deciding the right terms of address to be chosen. Generally speaking, in the case of close relationships where institutional power is irrelevant, it is age and gender that play the decisive role in deciding the right terms to choose. Women have low status against men and juniors low against seniors. However, in the case of outside relationship that involves institutional power, the institutional power has the final say. The strength of these factors in decreasing order are power, age and gender. Women use more polite forms, but old women receive polite forms from males, and even the oldest males use polite forms with people in authority.

4. DESCRIPTION OF ADDRESS FORMS IN CHINESE.

In Chinese, address terms could take the form of pronouns, personal names, kinship terms and other forms which I will call honorifics. In this part, I will explain the use of these four types of address forms and their functions.

4.1. Pronouns. Although Chinese has a system of pronouns, it was generally disfavored for possible address terms. In the old times, they were almost totally avoided. Since Confucian philosophy dictates that everything have a proper place by being assigned a correct name, therefore, a pronoun is not proper, because it is not a name itself but rather stands for a name. Use of a pronoun sounded rude and gave the impression of distance. Most of the time, honorifics were used in formal occasions and with strangers, while kinship terms were used for both close and distant relationships.

4.2. Personal Names. Although personal names, unlike pronouns, are real names and can be used as address terms, they were less often used than expected. The reason might be that they seldom indicate a person's position vis-à-vis his interlocutor and seldom tell about the relationship between speaker and addressee. Since Chinese tend to pay special attention to one's status vis-à-vis his interlocutor, terms that could bring out the relative relationship between the interlocutors are generally preferred. And what is more, it is impossible to address a stranger by his name. Again, most of the time, kinship terms were used for close relationships and honorifics for people with distance, and in some instances kinship terms have been extended for outside distant relationships.

4.3. Kinship Terms. One striking feature of the address system of Chinese is its extensive use of kinship terms for people outside the family as well as within the family. Kinship terms as address terms indicate intimacy, closeness and of course solidarity.

4.3.1. Kinship Terms for Close Relationships. The kinship system in Chinese is highly developed and elaborate. Several monographs have been published with the structure of kinship terms as the main focus (see Lin 1990 and Yang 1989). For the purpose of this paper, it serves as a starting point to say that the Chinese kinship system is characterized by strict patriarchy and strict order of seniority. Three of the five orders introduced by Confucius are about relationships in the family, namely, father/son, elder brother/younger brother and husband/wife. It reflects to a great extent the values that Chinese place on the family. The senior members have unchallenged high status and receive respect when they are addressed. Men generally are addressed with respect by women too. Finally, when a junior calls a senior, he has to choose a term not only appropriate in terms of the generation gap between them, but also appropriate in terms of the seniority of the addressee within his own generation.

4.3.2. Kinship Terms Extended to People Outside Family. Kinship terms have been extended to people outside the family circle and thus serve as general address terms, presumably to strike an accord with outsiders and adhere to the Chinese ideal of "a big family under one roof". As a result, neighbors and friends of the family are habitually addressed with kinship terms to signal solidarity. Oftentimes for the sake of showing politeness, these same kinship terms are used to strangers too. For example, "yi" is a term for one's mother's younger sister. It is now generally used for female friends of the family (of parents' age), and with the addition of an endearment morpheme it becomes the popular "a yi" (This term also has the advantage of saying the term without giving away the marital status of the addressee) (Hong 1985). The term has been extended for strangers too. Any female adult of one's parents' age could be called "a yi". Similarly, the use of many other kinship terms like "shu shu, ye ye, nai nai" and so on have also been extended.

5. HONORIFICS AND SELF-DEPRECATORY TERMS.

In addition to the above listed three categories of address terms, there is another category of terms of address. They are usually in pairs, denoting unequal relationship. The principle is that when one addresses others, s/he shows respects and elevates their status. When referring to oneself, one shows modesty and use self-deprecatory terms. They were used in unequal dyads, and to strangers. For example,

between officials and their subordinates

da ren (great man) xiao ren (little man)
Zun zhi (respectful office) bei zhi (humble office)

between the Emperor and court official

bi xia (your highness) chen (servant)

between husband and wife

Da ren (the great man) qie (mistress)

The use of honorifics was once very popular. One US diplomat to China in the late last century observed such an encounter (Young 1994:7)

What is your honorable cognomen?
The trifling name of your little brother is Wang.
What is your exalted longevity?
Very small. Only a miserable seventy years.
Where is your noble mansion?
The mud hovel in which I hide is in such or such a place.
How many precious sons have you?
Only so many stupid little pigs.

The above is only a rough translation, but the spirit is preserved. Interestingly, the emperor has a self-deprecatory term for himself—"zhen". Although it is supposed to show modesty, it is definitely not to be used by anyone else. The consequence would have been death penalty. By the way, it should be pointed out that most of these terms have fallen out of use since the founding of PRC.

6. FACTORS THAT DETERMINE THE CHOICE OF ADDRESS TERMS.

The following factors are presented with no intention of implying that they are independent of each other. Rather they are different dimensions of the same relationship.

6.1. Gender. Within the family, men and women did not have equal status. This was reflected in the address terms used by man and wife to each other although they were in a close relationship. In the past, in a typically authoritarian family, women called their husbands "da ren" (great man or master) or "dang jia de" (the one responsible for the family). They called themselves "jian ren" (undeserving person) or "qie" (mistress/concubine). Or to avoid these terms, they could call each other by invoking a third party, that is by referring to their kids, for example, "ta die" (children's father) or "ta niang" (children's mother).

6.2. Age. Age is directly associated with status, and seniors automatically receive respect when they are addressed. People from older generations are to be addressed with due respect, regardless of their gender. This is why, for example, a daughter-in-law

occupies one of the lowest positions in the family, but if she endured and became the mother and mother-in-law, she would enjoy prestige only next to her husband. Even people within the same generation are assigned strict rank order according to age. Siblings in a family can be called by their parents and by each other "lao da, lao er, lao san and so on" (the eldest, the second eldest, the third eldest, and so on). It is a popular form of address amongst gang members too. As a matter of fact, this practice is also very popular on college campus. Students coming from different parts of the country with diverse backgrounds and sharing the same dorm room figure out the rank of seniority and call each other accordingly as if they were brothers within a few days of meeting.

6.3. Power. By power, I mean the institutional power associated with one's official rank and position. Therefore it is usually connected with official settings. In these settings, institutional power has overwhelming weight in interpersonal communication, and Chinese generally show deference to authority by addressing those in power with due respect, regardless of age and gender (Pan 1995b).

7. CHANGE IN THE USE OF ADDRESS FORMS IN CHINESE.

So far, the discussion has mainly focused on the traditional use of address terms in Chinese, although some comparisons with present usage has also been discussed. Change started early this century with the end of thousands of years of feudalistic rule and the increasing influence from the West. But the fundamental change did not start until the founding of the People's Republic of China in 1949. Since then, many terms of address that denote inequality have gone out of use, for example "master, great man, little man, concubine", to name just a few, and address terms that denote socialist or communist ideals replaced those old terms of address (Ju 1991). Representing one of the ideals of communism, "tong zhi" became the most general term of address which denotes both equality and solidarity, used reciprocally practically between any two adults. Starting from early 1980s, the use of "tong zhi" began to lose its popularity and its use became more and more restricted. Scotton and Zhu (1983) document the change of "tong zhi" up to early 1980s. Now with nearly twenty years in between, we do not know for sure where "tong zhi" stands now. Although empirical study is not available, one interesting observation of the recent development of the term is that it is beginning to indicate distance as well as solidarity, depending on who is using the term to whom.

One immediate result of the unpopularity of "tong zhi" and its restricted use is that it leaves the space for other terms of address to fill in. So terms like "Mr., Mrs." etc, are coming back. But the reintroduction of these terms causes confusion to those who are accustomed to the old way of addressing. Look at the following example taken from Chen and Zhang (1986, 612). An American lady, Alice, who is 63, returns to China after her last visit thirty-five years ago when she was accompanied by Mr. Liu, now a retired high-ranking Communist party member. Experienced in dealing with foreigners, Mr. Liu greets Miss Alice the second time they meet with:

      Ai li si nu shi, jian dao ni fei chang gao xing.
      (Miss Alice, I am so glad to see you again.)

There nu shi means "lady". However, with less educated people, encounters with foreigners are not so. This time, Miss Alice comes to an apartment for a dinner, accompanied by Mr. Liu. Her hostess, laolao, is an old Chinese lady. They start a conversation over the dinner.

      Lao lao: wo shuo, zhei *wei* Ai tong zhi, *nin* jin nian duo da nian ji la?
      (Grandma: I say, this comrade Ai, how old are you?)
      Alice: liu shi san
      (Sixty-three)
      Lao lao: yao, zhen me shuo wo hai dei jiao *nin* lao jie jie ne.
      (Grandma: wow, in that case, I should have called you my old sister.)

This scene is funny because the grandma thought Westerners pronounce their surnames first as Chinese do. So she takes the first syllable of Miss Alice's given name (translated into Chinese as /ai-li-si/) for her surname. Also this is the first time that Grandma sits at the table with a foreigner, so she is being extremely polite by combining the polite form of second person pronoun "nin" with the honorific measure word for guest "wei". Not knowing the right term in Chinese for "Miss" or "lady", she opts for the most common one she knows, "tong zhi", for to her, this is a term of solidarity appropriate for a stranger. However, when she knows that her guest is sixty-three, she is much more confident about how to address her, and she uses "lao jie jie" (old sister). It is clear that she feels much more comfortable with the term and feels the use of the kinship term draws them closer immediately, so she sticks to the term for the rest of the encounter.

8. ADDRESSING AS INTERACTIONAL STRATEGY.

Appropriate relationships should be acknowledged and due respected should be paid by selecting the right address term. The above characterization may well give the impression that in Chinese society, a person's position relative to his interlocutor is to a large extent predetermined. It may seem that it follows that interlocutors do not have much leeway in selecting a term of address when interacting with each other. However, this does not need to be true. "Inappropriate" address forms do occur. But when "inappropriate" terms are chosen, it could be because the speaker has failed to recognize the requirements of the interaction, but more often than not, the choice itself is significant in that it is a deliberate choice for a specific interactional purpose. What follows illustrates how terms of address are chosen as interaction strategies.

8.1. Manipulative Use of Surnames. Surnames in Chinese are rarely used alone. Either full names are used or surnames are used together with some kind of title. The presence of surnames either shows distance or respect. So if a surname is used by itself, it carries

special meaning. The following line is taken from a fight scene of a play (Chen and Zhang 1986, 604).

> Xing wu de, yao shi ni na wo dang hou shua, ke bie guai wo fan hun.
> (You, the one with the surname of "wu", if you try to treat me as if I was dumb, then don't blame me if I should be not so polite with you.)

"Xing" is a verb which means "with the surname of" (like 'appeler' in French and 'heissen' in German), and "de" means "the one". The term of address "xing wu de" is intended as an insult or even a threat.

"Lao" and "xiao" can be used together with surnames in Chinese to signal solidarity, but the choice of "xiao" is marked in that while a person addressed by "xiao" plus surname is necessarily younger, one addressed by "lao" is not necessarily so. The following example from a movie adapted from a contemporary novel (Chen and Zhang 1986, 596) shows how this subtle difference could be manipulated.

Mr. Yuan, in his mid-forties, is the director of a state-owned enterprise. One day in early morning, on his way back home from the park where he practices shadowboxing, he runs into Mr. Li, a former classmate from twenty-two years ago. Mr. Li, though slightly younger than Mr. Gao, is already showing signs of age, as it is clear that he is losing his hair. Always self-conscious about his status as the head of the factory, Mr. Yuan greets his former classmate by saying:

> Hey, zhei bu shi xiao li zi ma? (Hey, isn't this "xiao li zi"?)

A guy in his forties and losing his hair certainly deserves "lao", but Mr. Yuan deliberately chooses "xiao li zi" instead. It looks like he is trying to be friendly, but in fact, he is trying to make clear to his interlocutor the power difference between them, as he is the director of the factory. For Chinese, one is more powerful in that one is older. Mr. Yuan not only chooses "xiao" instead, but also added a miniature morpheme "zi" to it. The net effect of the term is to signal difference in status.

"Lao shi", as professional title has been extended. However, if a faculty member approaches a janitor who is not regarded as a teacher of any kind and greets him by surname plus "lao shi", the janitor may be surprised. Using a term that the addressee apparently does not deserve has an obliging effect, and the janitor may think that the greeter has some kind of special favor to ask.

8.2. Manipulative/Selective Use of Kinship Terms. Technically, Chinese has terms for relationships that are five or more generations apart. However, for that purpose, a general word "zu zong" that means 'ancestors' is used. This is also the term used when one is addressing the family's ancestors at family sacrifices. Ancestors should be addressed with awe and in obedience. Therefore, one can imagine how desperate a frustrated parent is when he calls his rebellious and impossible child "wo de xiao zu zong" (my living ancestor).

Also, one can disparage people by calling them with a lesser term than expected or try to oblige others by calling them by an undeserved higher term. For example, one has a choice between "shu shu" and "bo bo" for his father's male friend, because both kinship terms are now extended for people outside the family. However, since "bo bo" literally means the older brother of one's father and "shu shu" the younger brother, the former is preferred when one wants to oblige the addressee for a favor.

Since many kinship terms have been extended to distant relationships or even to strangers, to stress closeness of a relationship, sometimes, these general terms which are less marked are disfavored against those that have not been extended and still serve to denote close relationship. For example, two families are close because the mother of family A and the father of family B are close friends or from the same town. Following the general rule, a male friend of one's family is expected to be addresses with the term "shu shu" (the extended use of the term for one's father's brother). However, the kids of family A might be instructed to call their family's male friend "jiu jiu" instead, a term reserved for their mother's brothers and which has not generally been extended. The effect is to emphasize the close tie of the two families. For the same reason, kids of family B might be told to call the family's female friend "gu gu", a term reserved for father's sisters.

8.3. Manipulative Use of Honorifics and Self-Deprecatory Terms. It is customary that one elevates one's addressee and lowers oneself when interacting with strangers or people in outside relationship. So one self-aggrandizing device for people of little importance is to call themselves "lao zi" ("the old one" which means "father"). Replying to his master's reprimand, a kneeling servant would be expected to say repeatedly:

> Xiao le bu gan. (me--the little one, dare not.)

But when angered, a desperate subordinate would probably shout out

> Lao zi cai bu chi ni zhei tao ne! (me—the old one, would not buy that stuff of yours!)

8.4. Selective Use of Variable Forms of Address. As documented by Scotton and Zhu (1983), the term "tong zhi" has lost favor and its use shows stratification according to age and occupation. This is to say that "tong zhi" has become a variable form.

As Kroger and Wood (1992) observe, in an analysis of address forms in former East Germany, there seems to be an inherent conflict between the demands for the expression of solidarity and for the acknowledgement of status in communist countries. The official ideology requires the expression of solidarity and the suppression of status, but the practical business of governing the nation requires the distribution of status as a reward for loyalty and as an incentive for extraordinary performance. This explains why in China during the time when "tong zhi" was the predominant form of address, a subordinate, when presenting a report for approval, chose to use the address "ju zhang tong zhi", a title plus "tong zhi" to achieve personal ends. The choice of the title "ju chang" (bureau head)

serves to show respect, but the addition of "tong zhi" mainly to signal solidarity, appealing to the superior to give the report favorable feedback.

One recent development in the use of the term is that it has taken on double meaning depending on who is using the term with whom. It can be observed that a subordinate uses it to narrow the distance between him and his superior. However, when the same "tong zhi" is used by the superior to the subordinate, it serves exactly the opposite function, that is, to distance the interlocutors. When a superior uses "tong zhi", he may imply that he is not talking to his subordinate as a person in a close relationship. What he does by calling the other party "tong zhi" is to invoke a third party who might as well be called "tong zhi". The implication is "a comrade by any other name will be the same". This is possible all because of the power difference between the interlocutors.

9.   ADDRESSING AS SOCIAL INDUCTION .

The above observation is mainly based on dyad interactions. There is another aspect of the use of address terms, that is, in triad interaction. As Blum (1997) observes, uttering relational terms in China has a unique performative power. She refers to the routine of introductions which consists of three participants. This usually involves the introduction of a junior to a senior by another senior. It is really like inducting the junior into the society, so I call it the ritual of social induction.

The following is a typical scene of social induction.

> (a parent): jiao yeye say grandpa
> the child: yeye grandpa
> grandpa: ahi, guai. Hao hai zi what a cute and obedient baby
> …… ……
> (the parent): gen yeye zai jian say goodbye to grandpa
> the child: yeye zai jian goodbye, grandpa

In this interaction of three parties, the animator of the words is the child, but the author of the words is the parent. All that the child has to do is to utter the words that acknowledge the status of the addressee being associated with age. It seems the parent, by teaching the child the right way to "talk" with seniors, is actually introducing the child into a society and helping him to find his own position in it. As being sure of one's position in the society is important, as preached by Confucius, this practice will go on until the child learns the correct way of addressing others, feeling secure about his own place in relation to others. Only then can we say the child has grown up and "dong shi le" (know things). One may point out that the child in the scenario does not know what he is saying. However, that does not matter. What really matters is the fact that he utters the term, and by uttering the term, he confirms the existing social hierarchy and the legitimacy of the relationship. The intention of the child when he uttered the term is less important than the actual uttering of the word and the mastering of a form through such practice. It is when the child refuses to say what the parent tells him to say that is really embarrassing for the parent. Then the parent would very likely mildly scold the child by characterizing the

child as "bu dong shi" (does not know things), and the grandpa would typically reply with "mei guan xi, ta hai xiao ne" (Don't take it too seriously. He is still young). By excusing the child of failing to say the right word of address, the senior at the same time also confirmed the subordinate status of the child—he is just a kid.

This kind of practice is part of a child's growing up. As he/she grows, other adult members of the family and close friends of the family will come along and continue to teach him and test him on the same lesson. Here is a typical play scene between an adult friend of the family and the child. Note the adult has a gift for the child.

> A: cai cai shu shu zhei you shen me? Guess what uncle has here?
> B: shi xiao ren shu! It is a cartoon book.
> A: xi huan ma? Xiang bu xiang yao? Do you like it? Don't you want it?
> B: xiang. Yes.
> A: xiang yao? Na jiao shu shu. Yes? Then call me "shu shu".

The child is encouraged to call the adult "shu shu", and there is an immediate reward associated with the uttering of the term. The lesson is clear: by uttering the term that by itself acknowledges the superiority of the addressee, the child acknowledges his own position in the social hierarchy. And to make the acknowledgment less painful, the society always makes it desirable in some way, because acknowledging one's inferior position also goes with social acceptance and reward. This kind of play almost always ends in the child uttering the terms the adult desired. And if any of the child's parents is present, he/she will insist that the child say Thanks. What is more, Thanks alone is not enough. Saying thanks without adding the address terms sounds very rude.

> Parent: Kuai xie xie shu shu. (Say Thank you, shu shu, quick.)
> The child: Xie xie shu shu (thank you, shu shu.)
> Adult friend: ai guai, bu yong xie. (what an obedient kid. No need to say
> thanks.)

It seems that China is a society which attaches great importance to preserving the existing social hierarchy. Young people are taught whenever the opportunity arises to confirm the legitimacy of the social structure and are encouraged to enter the social network accompanied by a rewarding experience.

Still another case where the ritual of induction is observed is when the would-be daughter-in-law is introduced into the family for the first time. Blum (1997) is right in observing that one key step to being inducted into the family is by learning the right terms of address for each member of the family. But again, I would add that this ritual is also performed with reward. While it is not easy to call people who up to then were virtual strangers by intimate kinship terms, the efforts are always rewarded with some material gain. For example, when the would-be daughter-in-law is expected to call the would-be father- and mother-in-law as "mom" and "dad" for the first time, the parents should have money ready wrapped in a red envelope. This is called "kai kou qian", which means "fees for opening the mouth". Neighbors will judge whether the would-be

daughter-in-law is a member of the family by asking the would-be mother-in-law "Did she call you mom yet?"

10. ADDRESSING AS NEGOTIATION.

Interlocutors always have the choice of different terms and the choice itself is often socially significant. That an interlocutor has chosen an unexpected address term to control the interaction does not mean the other is obliged to accept it no matter what interactional power is at stake. For example, when someone feels obliged by being addressed with a term he does not deserve, he may deny it and opt for a more appropriate one.

> Parent: jiao Wang bo bo. (say Wang bo bo.)
> Child: Wang bo bo. (Wang bo bo.)
> Wang bo bo: bie bie, bu gan dang, jiao shu shu hao le. (please don't. I dare not take it.
>   Just call me shu shu.)

The addressee feels the term of "bo bo" is undeserved because he is not obviously older than the child's father, also because he suspects that his interlocutor might have some favor to ask of him. In addition, this encounter again supports the idea that the child is not the author of his speech. The addressee's answer "bu gan dang" is not directed to the child but rather to the parent.

11. CONCLUSIONS.

Among the four types of address forms discussed above, kinship terms are the most widely used, for both close relationship and distant relationships. Age, gender and institutional power in ascending order of weight are the key factors influencing the choice of address terms. Although address forms index such macro-social realities such as age, gender, solidarity and status, choice of address terms are often strategic in that they are consciously manipulated by interlocutors, either to negotiate social roles and relationships or to redefine the terms of ongoing interaction.

REFERENCES

BLUM, S.D. 1997. Naming Practice and the Power of Words in China. *Language in Society* 26.357-79.

BROWN, R. 1965. *Social Psychology*. Chicago: Free Press.

CHEN, KAI-YAN and ZHANG, LING. 1986. *Anthology of Chinese Movies*. Beijing: Chinese Movies Publisher.

ELLIS, R. 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

GU, YUE-GUO. 1990. Politeness Phenomena in Modern Chinese. *Journal of Pragmatics* 14(2).237-58.

HONG, B. 1985. Politeness in Chinese: Impersonal Pronouns and Personal Greetings. *Anthropological Linguistics* 27(2).204-13.

IDE, S., IKUTA, S., KAWASAKI, A., and OGINA, T. 1992. The Concept of Politeness: An Empirical Study of American English and Japanese. Chapter in *Politeness in Language: Studies in Its History, Theory and Practice*, ed. by Watts, R. J., Ide, S., and Ehlich, K. New York: Mouton de Gruyter.

JU, ZHUCHENG. 1991. The Depreciation and Appreciation of Some Address Terms in China. *Language in Society* 20.387-90.

KROGER, R. and WOOD, L. 1992. Are the Rules of Address Universal? *Journal of Cross-Cultural Psychology* 23(2).148-62.

LIN, MEI-JUNG. 1990. *Han yu chin shu cheng wei ti chieh kou fen hsi* (Analysis of Chinese kinship terms). Taipei: Tao Hsiang Publisher.

PAN, YULING. 1995a. Addressee, Setting, and Verbal Behavior: How Relevant are They in Foreign Language Teaching? *Georgetown University Round Table on Languages and Linguistics 1995*. Washington, DC: Georgetown University Press.

PAN, YULING. 1995b. Power Behind Linguistic Behavior: An Analysis of Politeness Phenomena in Chinese Official Settings. *Journal of Language and Social Psychology* 14(4).434-61.

SCOLLON, R and SCOLLON, S. 1991. Topic Confusion in English-Asian Discourse. *World Englishes* 10(2):113-25.

SCOTTON, C. M. and ZHU, WANJIN. 1983. Tongzhi in China: Language Change and its Conversational Consequences. *Language in Society* 12:477-94.

YANG, YING-CHIN. 1989. *Ku chin cheng wei tzu tien* (Dictionary of Chinese kinship terms). Ho-fei: Huang-shan shu she.

YOUNG, L.W. 1994. *Crosstalk and Culture in Sino-American Communication.* Cambridge: Cambridge University Press.

SEMANTIC-DISCOURSE ACCOUNT FOR THE DATIVE-SUBJECT CONSTRUCTION IN JAPANESE

KYOKO MASUDA

University of Arizona

1. INTRODUCTION.

Traditionally, Japanese linguists have proposed that postpositional particles are used to mark case relations in Japanese sentences. They include NOMINATIVE *ga,* ACCUSATIVE *o,* DATIVE *ni*, and GENITIVE *no* as shown in (1).[1]

    (1)   Taroo ga   Hanako no   sensei   ni   denwa o shita.
                 Taro NOM   Hanako GEN   a teacher DAT call   ACC did
                 'Taro made a call to Hanako's teacher'.

However, some Japanese verbs of possession, capability, or perception and verbs with an ability auxiliary show variation in their patterns [Noun phrase *ga*, Noun phrase *ga* Verb phrase] or an [Noun phrase *ni*, Noun phrase *ga*, Verb phrase] as in sentences (2) - (5).

    (2)   Taroo ga/ni   oniisan   ga   iru.
                 Taro   NOM/DAT   brother   NOM have
                 'Taro has brothers'.   [possession]

    (3)   Taroo ga/ni furansugo   ga   wakaru.
                 Taro   NOM/DAT French   NOM   understand
                 'Taro understands French'.   [capability]

    (4)   Taroo ga/ni furansugo   ga   hanas-eru.
                 Taro   NOM/DAT French   NOM   Speak-can
                 'Taro can speak French'.   [verb with an ability auxiliary]

    (5)   Taroo ga/ni   nagareboshi ga   mie-ta.
                 Taro   NOM/DAT   shooting star NOM   can see-PAST   [perception]
                 'Taro could see a shooting star'.

This is called the dative-subject construction or the *ga/ni* alternation, which has been examined in two different approaches: syntactic-based studies (Kuno 1972, Shibatani 1977, 1978, Saito 1982, Takezawa 1987, Miyagawa 1989, Dubinsky 1992, Sadakane and Koizumi 1995) and non-syntactic-based studies (McGloin 1980, Watanabe 1984, Kabata 1998). In the previous literature, this alternation was claimed to be a syntactic matter; for instance, Kuno (1972) considered it as the result of an avoidance of *ga-ga* repetition, while Shibatani's (1977, 1978) analysis demonstrated that it is the syntactic requirement of at least one *ga*-marking in a non-embedded sentence. Recently, however, Kabata

(1998) suggested that it is a semantic/cognitive phenomenon reflecting the nature of agent/experience continuum. The purpose of this paper is to show that not only semantic/cognitive factors but also one of the discourse factors, information status, is crucial to determining the selection of *ga/ni*. I will focus on the Japanese nominative marker *ga* as discussed in Kuno (1972) and Chafe (1970, 1976, 1987). To support my claim, I will provide introspective evidence from 20 native Japanese speakers who were tested on the acceptability of *ga/ni* markings.

1.1. Previous Study: Kabata (1998). Kabata (1998), argues that the Japanese case alternation is not simply a syntactic matter, but that it is more fully explained by Langacker (1991a, 1991b)'s archetypal model. According to Langacker (1991a: 238-9), the archetypal AGENT is a person who volitionally initiates a physical activity, while the archetypal EXPERIENCER is characterized as a sentient entity engaged in a mental activity. The model that Kabata applies to Japanese is illustrated in Figure 1 taken from Kabata and Rice (1997: 111).

FIGURE 1. KABATA AND RICE (1997) ARCHETYPAL MODEL
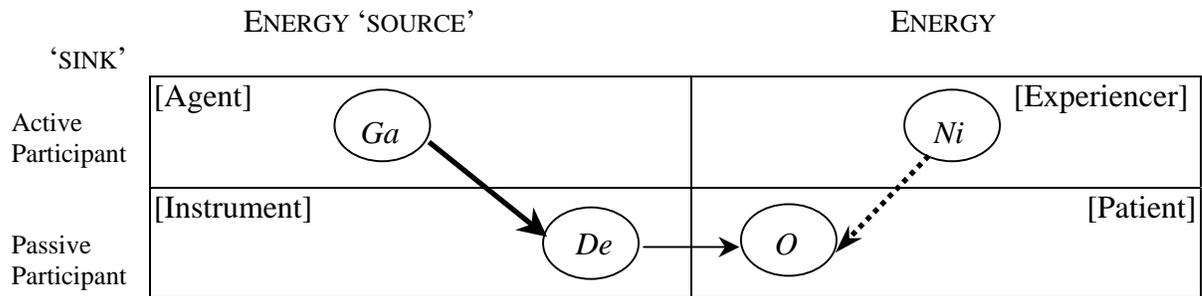
Describing the interaction of clause participants using a metaphor of ENERGY TRANSMISSION, this model captures the essential non-categorical nature of agentive and experiencer roles. As shown in Figure 1, in Japanese *ga* represents the highest rank in the energy flow (ENERGY-SOURCE), coding both volitionality and agentivity in active participants, while *ni* illustrates the lowest rank in the energy (ENERGY-SINK), coding a lack of volitionality and agentivity in active participants.

Kabata claims that several semantic factors have an impact on the selection of *ga/ni* such as the volitional act, the inchoative act, and negation with an intensifier. Notice that these semantic factors are closely related to transitivity components in discourse that Hopper and Thompson (1980) claim to be a central role in understanding of clause structure in human language, e.g. participants, kinesis, volitionality, affirmation, agency, and affectedness, among others. The sentences (6) – (10) from Kabata (1998), along with her grammaticality judgments, show how *ga/ni* are used in the Japanese dative construction according to semantic factors.[2]

(6)   Taroo ga/ni   Masako no kimoti   o/ga   wakaru.
            Taro   NOM/DAT   Masako GEN feeling   ACC/NOM   understand
           'Taro understands Masako's feeling(s)'.

(7)   Taroo ga/*ni   Masako no   kimoti o/*ga   wakarooto-si-ta.
            Taro NOM/DAT   Masako GEN   feeling ACC/NOM
           understand:VOL-do PAST
           'Taro tried to understand Masako's feeling(s)'.

(8)   Taroo ?ga/ni   Masako no   kimoti *o/ga   wakatte-ki-ta.
            Taro   NOM/DAT   Masako GEN   feeling ACC/NOM   understand-
                 come-PAST
           'Taro came to understand Masako's feeling(s).'

(9)   Taroo *ga/ni(wa)   Masako no   kimoti *o/ga   wakara-nai.
            Taro NOM/DAT(TOP)   Masako GEN   feeling ACC/NOM   understand-NEG
           'Taro does not understand Masako's feeling(s)'.

Sentence (6) is neutral while sentence (7) contains a volitional act, *-oo tosita* (tried to). These sentences suggest that *Taroo* in the first noun in sentence (7) is more likely to be marked with *ga* when it functions as an energy-source or high transitive referent. However, if a sentence has expressions showing the inchoative act or negation as shown in (8) and (9), respectively, then the first noun is more likely to be marked with *ni*, indicating an energy-sinking or low transitive referent. Sentence (8) is an example for an inchoative expression, *wakatte kita* (came to understand), hence requiring *ni*, not *ga*. Finally, sentence (9) has the negation *-nai* where *ni* is followed by an optional particle *wa*. Based upon the observation that *ga* marks volitional participants in affirmative contexts like sentence (7), and that *ni* marks non-volitional participants in negative contexts like sentence (9), it follows that *ga* is expected to be high transitivity while *ni* is low transitivity.

    Kabata conducted a survey with 112 Japanese high school students whom she asked to fill in the blanks by choosing a particle in clauses with different semantic factors. An example from her survey is illustrated in sentence (10). Table 1 summarizes the main findings.

(10) Hanako ----- asa   hayaku   okireru.
      Hanako   morning early   can get up
      'Hanako can get up early in the morning'.

      1. *ga*   2. *niwa*   3.*ni*   4.others

One of the findings is that semantic factors such as volitionality, negation with an intensifier, and inchoativity crucially influence the selection of particles. As Table 1 shows, *ga* was more acceptable in combination with a high volitinality act, while *ni* was most acceptable in combination with a negation act and with an inchoative act. In spite of the strong findings in her study, some aspects remain to be reexamined.

TABLE 1. KABATA 1998: FIVE TYPES OF SEMANTIC MODIFICATIONS WITH
THE VERB *WAKARU* (understand)

Stimulus Sentence	Condition	*ni*	*ga*	others
a) Watasi ----- kareno kimoti (o/ga) wakatta. 'I understood his feelings.'	NEUTRAL	79	3	27
b) Watasi ----- kareno kimoti (o/ga) wakarootoshita. 'I tried to understand his feelings.'	VOLITIONAL	2	25	82
c) Watasi ----- kareno kimoti (o/ga) wakattekita. 'I came to understand his feelings.'	INCOHATIVE	66	2	41
d) Watasi ----- kareno kimoti (o/ga) wakaranakatta. 'I did not understand his feelings.'	NEGATIVE	90	1	18
e) Watasi ----- kareno kimoti (o/ga) doositemo wakaranakatta. 'I did not understand his feelings by any means.'	NEGATIVE/ INTENSIFER	100	0	9

*The number in Table 1 mean raw score for 5 different types of semantic modifications with the verb *wakaru* (understand). The totals reflect the number of subjects who chose the particular particles.

    First, in Kabata's survey study there is a discrepancy between her claim and experimental materials that she used illustrated in (11) below.

        (11) Hanako ---   asa hayaku okireru.

        Example   1. *ga*   2. *niwa*   3.*ni*   4.others

        Analysis   *ga*   vs.   *ni*   vs.   others
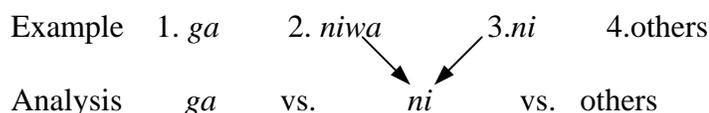
She included the choice of *niwa*, in which *ni* is followed by a topic marker *wa*. When she analyzed the data, she combined *niwa* and *ni*, since she assumes that *ni* and *niwa* are used interchangeably, and therefore semantically the same, although she acknowledges that *niwa* adds some discourse-level function.   However, I found this analysis to be problematic as there might be a high possibility that the choice of *niwa* increases the high frequency of *ni* selection. In her research design, we cannot tell to what extent a choice of *niwa* increased the high frequency of *ni* because she put *ni* and *niwa* together in her analysis.
    Secondly, Kabata's findings are limited by her choice to exclude discourse factors, which is closely related to my first point. Namely, in her survey study she did not examine the discourse factor of information status. Further the previous research ignores the inherent relationship between discourse function and the semantic nature of *ga.*[3] This paper will attempt to examine both semantic and discourse factors of *ga*.

2. Methodology

A survey study was conducted to investigate the manner in which the acceptability of *ga* and *ni* marking can be attributed to semantic/discourse factors discussed in the previous section. Twenty adult Japanese subjects participated in a survey study. They have been living in the U.S for a period of four months to eight years. The average age was about thirty-two years old. All participants answered that they speak Japanese in their daily life and participated on a voluntary basis.

Participants were asked to fill in blanks in the stimuli sentences with either *ga, ni, niwa*, or *de*.[4] The topic marker *wa* was not included in the choice, but they were instructed to use other particles if they needed to. The sentences were presented in a randomized order with a distracter, a calculation problem.

In the survey there were 72 experimental sentences written in Japanese which varied in terms of several semantic dimensions. Semantic conditions which are relevant are the following: 1) four semantic cues such as neutral, volitional, negative with intensifier, and inchoative, and 2) three different types of subject nouns which will be discussed later in this section.

Unlike Kabata, I controlled the particles of the object markers since in my first two pilot studies, the subjects gave relatively consistent answers.[5] My study is almost identical to Kabata's study except for two points: 1) adding the new dimension of information status in subjects, and 2) analyzing data by separating the responses of *ni* and *niwa*.

This paper will adapt Chafe's (1987) notion of new/old information in discourse. He maintains that given (or old) information means the knowledge which a speaker assumes to be in the consciousness of an addressee, while new information means what a speaker assumes s/he is introducing into an addressee's consciousness by his/her utterance. Based upon Chafe's notion, I use three types of subjects: *aru otoko* (certain man), *sono hito* (that person), and *watashi* (I), in order to examine the interaction between particle choice and information status of the subject, i.e. new information, given information from the previous discourse, and given information from the conversational context. These are shown in sentences (12) – (14).

(12)   Aru otoko (*ni*, *ga*, *niwa*, *de*)   A-maunten   made   hashirouto-doryokushi-ta.
        Certain man   A-mountain   to   run: VOL-try-PAST
        'Certain man tried to run to A-mountain'.

(13)   Sono hito (*ni*, *ga*, *niwa*, *de*) furansugo   *ga*   wakaru.
        That person   French   NOM   understand
        'That person understands French'.

(14)   Watashi (*ni*, *ga*, *niwa*, *de*)   nagareboshi   *ga*   mie-ta.
        I   shooting star NOM   can see-PAST
        'I could see a shooting star'.

Aru otoko (certain man) in (12) is a good example of new information that Chafe discusses. *Aru* (that) is typically used when a speaker introduces a character for the first time, or when a speaker establishes a certain setting in discourse. *Sono hito* (that person) in (13) is an example of the given information which has already been introduced in discourse. *Watashi* (I) in (14) is an example of what Chafe (1987) analyzes as given information, since first and second person referents such as 'I' and 'you' acquire the given status naturally from the conversational context itself.

3. RESULTS AND DISCUSSION.

3.1. SEMANTIC CONDITIONS.

There are two conditions that can interact with particle choice in the Japanese dative construction: semantic conditions and reference types. First, Table 2 (Interaction between Semantic Conditions and Particle Choice) which was obtained based upon the 180 responses, shows how semantic conditions interact with particle choice in the Japanese dative construction.

TABLE 2. INTERACTION BETWEEN SEMANTIC CONDITIONS AND PARTICLE CHOICES[6]

Conditions	*ga*	*ni*	*niwa*	*wa*	Total
Neutral	23	23	108	33	187
	12.2%	12.2%	57.8%	17.8%	100%
Volitional	91	3	15	72	181
	50.3%	1.7%	8.3%	39.8%	100%
Negation with intensifier	10	1	159	21	191
	5.3%	0.5%	83.3%	11%	100%
Inchoative	43	23	58	58	182
	23.6%	12.6%	31.9%	31.9%	100%

As can be seen, overall differences in particle choices were found depending on different semantic factors such as volitionality, neagativity with intensifiers, and inchoativity. While in neutral patterns the most frequently used particle was *niwa* (58%), other particles were in the range of between 12% and 18%. Notice that in all other semantic conditions *niwa* was more frequently choosen than *ni*, although in the volitional condition the number may not indicate this is a strong case. This is not observable in Kabata's survey since *ni* and *niwa* were classified into one group in her analysis.

As shown in Table 2, however, in volitional sentences the frequency of *ga* was 50%, while the frequencies of *ni* and *niwa* were only 2% and 8%, respectively. This finding is in accordance with the earlier discussion of *ga* as expressing high-energy source or high transitivity referent in volitional sentences which often have a high degree of transitivity.

It is also important to note that in negative sentences with an intensifier, the most frequently selected particle was *niwa* (83%), whereas other particles were rarely used except for *wa* (11%). This result suggests that the negative takes not only an energy-sink or low transitivity particle like *ni*, but also takes *wa*, which has a strong connection with negation. This finding is consistent with McGloin (1980) who emphasizes that the NP *ni wa* in Japanese dative construction occurs more easily with negation[7].

In contrast, as described in Table 2 in inchoative sentences, a somewhat different pattern was observed. Surprisingly, in this study the frequency of *ga* (24%) was higher than that of *ni* (13%) and lower than that of *niwa* (32%). This is one discrepancy between Kabata's study and this study, since in the former extremely low frequency of *ga* was observed in inchoative sentences (approximately 2 %) and quite high frequency of *ni/niwa* (approximately 61%) was found. These different results seem to suggest that whether or not inchoative sentences have low transitivity is unclear; therefore, there is a need to re-examine this issue[8]. These findings confirm Kabata's study that the presence of volitionality increased the acceptability of *ga,* while the presence of negation with an intensifier failed to increase the *ga* response since negation has a property of energy-sink or low transitivity. It is also found that in the case of negative with intensifier, the frequency of *niwa* (83.3%), not *ni*, increased. This suggests that high frequency of *ni* in Kabata's analysis is due to the frequency choice of *niwa*, rather than *ni* alone.

3.2. REFERENCE TYPES.

Table 3 (Interaction between Reference Type and Particle Choices) illustrates a different tendency in particle selection among reference types. The percentage of each type of noun was calculated from 430 responses.[9]

TABLE 3. INTERACTION BETWEEN REFERENCE TYPE AND PARTICLE CHOICES

	ga	ni	niwa	wa	total
Watashi (I)	79	19	223	112	433
	18.1%	4.8%	51.6%	25.9%	
Sono hito (that person)	78	30	215	107	430
	18.1%	6.9%	50%	24.9%	
Aru otoko (certainman)	148	34	158	98	438
	33.8%	7.8%	36%	22.4%	

* No one answered *de*; therefore, no category for this particle is provided.

As can be seen with *sono hito* (that person) and *watashi* (I), both old information pronouns shared a similar pattern; *niwa* was the most frequently used (approximately 50%), while *ga* was not highly preferred in both references. What is striking is the increase in preference for *ga* when *aru otoko* (certain man) was used as a subject.

Compared to the other types of pronouns such as *sono hito* (that person) and *watashi* (I), it is evident that *ga* is the most preferred particle for *aru otoko* which carries new information status. This finding suggests that there is a need to consider the information status of the subject as an influence on particle selection. A promising interpretation of this finding lies in discourse analysis where the distinction between given/new information is central. Namely, as a word, *aru otoko* carries new information, and as new information is marked with *ga* in Japanese, it follows that *aru otoko* is marked with *ga* rather than *ni.*

3.3. INTERACTIONS BETWEEN SEMANTIC AND DISCOURSE FACTORS.

It has been shown that both semantic factors such as volitionality, negation with intensifiers, and inchoativity affect the selection of *ga/ni.* We have also observed that the information status affects the selection of *ga/ni.* However, what is not clear at this point is whether there is any interaction between semantic factors and the information status. There are cases suggesting that two factors, a semantic factor (an energy transmission or transitivity factor) and a discourse factor (new/old information status), are competing with or reinforce each other as in sentences (15) - (17).

Competing factors

(15)   Aru otoko (*ga* 3, *niwa* 15, *wa* 2)   nagareboshi   *ga*   dooshitemo
        mienaka-tta.
        A man   shooting star   NOM   by any means   can not
        see-PAST
        'A man could not see a shooting star by any means'.

Reinforcing factors

(16)   Watashi (*niwa* 20)   nagareboshi   *ga*   dooshitemo   mienaka-tta.
        I   shooting star   NOM   by any means   can not see-PAST
        'I could not see a shooting star by any means'.

(17)   Sono hito (*niwa* 20, *wa*1)   nagareboshi *ga*   dooshitemo   mienaka-tta.[10]
        That person   shooting star NOM   by any means   can not see-
                    PAST
        'That person could not see a shooting star by any means'.

In sentence (15), the first noun *aru otoko* (certain man) carries the new information status calling for *ga*, but the semantic condition expressed by *doushitemo mienakatta* (could not see by any means) indicating low transitivity due to negative with intensifier, *doushitemo – nai* (not… by any means) calls for *niwa*. As a result, the discourse factor of new information status and the semantic factor of low transitivity appear to be competing with each other in this sentence. The semantic factor seems to be stronger as fifteen out of

twenty participants chose *niwa*, while only three chose *ga*. An intriguing finding can be observed by comparing this to sentence (16) and (17); in sentence (16), where *watashi* (I) indicates given information, no one answered *ga* which implies that semantic and discourse factors are not competing with, rather reinforcing each other, which resulted in no occurrence of *ga*. This seems to hold true in (17) as *sono hito* (that person) also carries given information status is used.

Table 4. Interaction between Semantic Factor and Discourse Factor

	High Transitivity (volitional) Energy-Source					Low Transitivity (Negation w/ Intensifier) Energy-Sink				
	ga	*ni*	*niwa*	*wa*	total	*ga*	*ni*	*niwa*	*wa*	total
New Information *aru otoko* (certain man)	19	0	0	17	83	19	7	75	40	168
	80%	0%	0%	20%	100%	28%	4%	45%	26%	100%
Old Information *sono hito* (that person), *watashi* (I)	70	3	16	54	161	5	18	228	61	326
	43%	2%	10%	45%	100%	6%	5%	70%	19%	100%

* No one answered *de*; therefore, no category for this particle is provided.
* Totals are different as the numbers of sentences that fall into each category are different.

Table 4 (Interaction between Semantic Factor and Discourse Factor) summarizes the results of interactions between a semantic factor (energy transmission or transitivity) and a discourse factor (given/new information status). In this table, volitionality and negative with intensifiers were chosen as factors representing high/low transitivity, respectively.

First, this table clearly shows that in a clause that contains a combination of new information and high transitivity, *ga* was strongly preferred (80%). Secondly, in a clause which includes both given information status such as *watashi* and *sono hito*, and negation with an intensifier (energy-sink or low transitivity), *niwa* was the most preferred particle to be used (70%). Lastly, the results indicate that the use of *aru otoko* (certain man) carrying new information significantly increased the use of *ga* in each of the high/low transitivity clauses. Within the energy-sink or the low transitivity context, the usage of *aru otoiko* (certain man) increased the frequency of *ga* from 6% to 28%, while in the energy-source or the high transitivity context, it increased the frequency of *ga* from 43% to 80%. It becomes clear that the discourse factor of given/new information status in relation with semantic factors plays a role in affecting the choice of *ga/ni* in Japanese dative-subject construction.

4. CONCLUSION.

This survey study has provided supporting evidence for Kabata's (1998) claim: *ga/ni* alternation is not just a syntactic matter, but is cognitively motivated by several semantic cues such as negation with intensifier, volitional, and inchoative in a sentence. This survey has also offered a piece of evidence suggesting that not only the semantic cues

mentioned, but also the old/new information status of subjects plays a crucial role in the selection of *ga* or *ni* as subject markers. In the future, other devices such as natural data including written and oral discourse should be investigated with respect of the issues raised in this survey.

REFERENCES

CHAFE, WALLACE. 1970. *Meaning and the Structure of Language*. Chicago: University of Chicago Press.

CHAFE, WALLACE. 1976. Giveness, Contrastiveness, Definitiveness, Subjects, Topics, and Point of View. *Subject and Topic*, ed. by Charles N. Li, 27-55. New York: Academic Press.

CHAFE, WALLACE. 1987. Cognitive Constraints on Information Flow. *Coherence and Grounding in Discourse,* ed. by R. S. Tomlin, 23-51. Amsterdam: John Benjamin.

DUBINSKY, STANLEY. 1992. Case assignment to VP-Adjoined Positions: Nominative Objects in Japanese. *Linguistics* 30.873-910.

HOPPER, PAUL J. and THOMPSON, SANDRA, A. 1980. Transitivity in Grammar and Discourse. *Language* 56. 251-299.

KABATA, KAORI. 1998. On the Agent/Experience Continuum: A Cognitive Approach to the *ga/ni* Alternation in Japanese. Unpublished manual.

KABATA, KAORI AND SALLY RICE. 1997. Japanese ni: The Particulars of a Somewhat Contradictory Particle. *Lexical and Syntactical Constructions and the Construction of Meaning*, ed. by Marjolijn, Verspoor, Kee Dong Lee and Eve Sweetser, 107-127. Amsterdam: John Benjamins.

KUNO, SUSUMU. 1972. Functional Sentence Perspective: A Case Study from Japanese and English.
Linguistic Inquiry 3.269-320.

KURODA, SHIGEYUKI. 1965. *Generative Grammatical Studies in the Japanese Language*. Ph.D. Dissertation. MIT.

MAYNARD, K. SENKO. 1987. Thematization as a Strategies Device in the Japanese Narrative. *Perspectives on Topicalization: The Case of Japanese wa*, ed. by John, Hinds, Senko K. Maynard, and Shoichi Iwasaki, 57-82. Amsterdam: John Benjamins.

LANGACKER, RONALD W. 1991a. *Foundations of Cognitive Grammar*, Vol. 2. Stanford: Stanford University Press.

LANGACKER, RONALD W. 1991b. *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. Berlin: Mouton De Gruyer.

MCGLOIN, HANOKA. N. 1980. *Ga/ni* Conversion Re-examined. *Papers in Japanese Linguistics* 7.65-77.

MIYAGAWA, SHIGERU. 1989. Syntax and Semantics 22. Structural and case marking in Japanese. New York: Academic Press.

ONO, TSUYOSHI, THOMPSON, SANDRA, A. and SUZUKI RYOKO. 2000. The Pragmatic Nature of the So-Called Subject Marker *ga* in Japanese: Evidence form Conversation. *Discourse Studies* 2.35-64.

SADAKANE, KUMI and KOIZUMI, MASATOSHI. 1995. On the nature of the 'dative' particle *ni* in Japanese. *Linguistics* 33.5-33.

SAITO, MAMORU. 1982. Case Marking in Japanese: A Preliminary Study. Unpublished Manuscript. MIT.

SHIBATANI, MASAYOSHI. 1977. Grammatical Relations and Surface Cases. *Language* 53.789-809.

SHIBATANI, MASAYOSHIO. 1978. Nihongo no bunseki. Tokyo: Taisyukan.

TAKEZAWA, KOICHI. 1987. A Configurational Approach to Case-Marking in Japanese. Ph.D. Dissertation. University of Washington.

WATANABE, YASUKO. 1984. Transitivity and Evidentiality in Japanese. *Studies in Language* 8(2).235-251.

NOTES

1. The abbreviations used for the gloss in this paper are as follows:

NOM	nominative	ACC	accusative
DAT	dative	TOP	topicalizer
GEN	genitive	PAST	past tense
NEG	negative	CONJ	conjuncitve particle
VOL	volitionality		

2. I have modified Kabata's glossing slightly for the sake of clarity.

3. The litrature on the Japanese subject marker *ga* has been discussed by contrasting it with the so-called topic marker *wa*. For instance, Kuno (1972) argues that *ga* is used when new and unpredictable information is presented. Chafe (1970) states that "in Japanese…the surface structure particles *wa* reflecting old information, and *ga*, reflecting new." There are other studies that share the similar view of *ga* (Chafe 1987, Maynard 1987, Ono, Thompson and Suzuki 1999). It is generally accepted that in the subject position of matrix sentences, *ga* is used for new information.

4. The Japanese postposition *de* is an instrument marker and translated as 'by means of' and 'by' etc. *De* is different from other choices in this survey in that it cannot be used to mark the dative subject in Japanese.

5. In my first two pilot studies, I found that all five subjects answered *ga* for an object marker in sentence (i), and four answered *ga* and only one person answered *o* in the sentence (ii). I used *ga* for an object marker in my survey because *o* was less likely to be used.
    (i) Watashi (   ) furansugo (   ) hanaseru. (I can speak French.)
    (ii) Kare (   ) furansugo (   ) hanaseru. (He can speak French.)

6. Since no participant chose *de* for the dative subject marker, no category for this is provided in Table 2. Since some participants answered *wa* on a voluntarily basis, the total numbers for each semantic condition were slightly different.

7. In her paper, McGloin cites Kuroda (1965) who claims that negative sentences are uttered in a context where corresponding affirmatives have already said or the speaker assumes that the hearer knows. Likewise, the Japanese dative construction calls for a context where the predication is discussed. According to McGloin, it is reasonable to have a link between a *niwa*-marked noun phrase and negation in that both need some kind of presumption in context.

8. Kabata states that inchoative is in a context 'biased any from willful effort' which I

interpret to mean that she assumes that inchoative is likely to be marked by *ni*, an energy-sink marker.

9.   The total number of particles in Table 3 (the information status of the subject) is different from the one in Table 2 (the semantic factors). This is because I did not include data from the semantic conditions due to peculiarity of combining all semantic conditions discussed in this study. I specified three relevant cues such as *kimoti-wakaru* 'understand one's feeling', *furansugo-wakaru* (understand French), and *furansugo-hanaseru* 'can speak French', since all three cues allow me to test all three semantic factors, volitional, negative, and inchoative.

10.   The topic marker *wa* was not included as a choice, but they were instructed that if they needed other particles, they could write down the particles. The fact that 20 participants answered *niwa* and one participant answered *wa* means that there was one participant who answered that both *niwa* and *wa* were equally appropriate.

THE *GO-AND-VERB* CONSTRUCTION IN A CROSS-LINGUISTIC PERSPECTIVE: IMAGE-SCHEMA BLENDING AND THE CONSTRUAL OF EVENTS[*]

Anatol Stefanowitsch

Rice University

1. INTRODUCTION.

This paper investigates a construction in English (and in some other languages, including the Scandinavian and the Ibero-Romance languages), which I will refer to as the *go-and-Verb* construction. Examples from English are *Look what you've GONE AND DONE!* or *Nobody believed she would climb Everest, but she WENT AND DID it.*

Most of the literature on this construction (referred to as 'pseudo-coordination' (e.g. Joseffson 1991), 'fake coordination' (e.g. Carden and Pesetzky 1979), or 'hendiadys' (e.g. Poutsma 1928) deals almost exclusively with its formal properties. In contrast, my paper focuses on its semantics. Section 2 discusses the major uses of this construction in English and compares them to those in other languages. It will become clear that the semantics of the construction seem quite disparate at first glance: even in the two examples given above it is used to express such seemingly unrelated notions as 'disapproval' and 'unexpectedness.'

I then show that it is possible to offer a unified account of those differing uses. Section 3 argues that all uses of the *go-and-Verb* construction are motivated by a combination of the image schematic properties of the verb *go* and the more richly specified semantics of whichever second verb occurs in a particular expression. Section 4 discusses the precise nature of this mechanism: the image-schematic properties evoked by *go* are blended (or 'fused') with the event structure of the second verb to allow the speaker to construe the event denoted by the second verb in accordance with the image-schematic meaning of *go* (cf. Ekberg 1993 for a similar approach).

2. THE DATA.

In this section I will present and discuss the data, but before I do so, some terminological issues must be clarified.

First, the construction discussed in this paper is referred to as the '*go-and-Verb* construction,' regardless of the language under discussion at any particular point. This is not meant to suggest that we are dealing with one and the same construction across different languages (I will return to this point). Likewise, verbs corresponding to English

go are referred to as 'go.' When I refer to a verb from a specific language, this is indicated.

Second, the other verb in the *go-and-Verb* construction is referred to as the *main verb*. This term reflects the intuitive notion that the second verb in this construction carries most of the semantics, while *go* functions somewhat like an auxiliary; however, I will not argue this latter point, so the term 'main verb' may simply be regarded as a useful label.

The syntactic properties of the *go-and-Verb* construction will not be discussed in any detail in this paper, but there are two facts that will be relevant later. First, there is the fact that the construction involves two verbs which are coordinated by *and* (rather than by simple juxtaposition, as in *Go jump in a lake*). Second, syntactic evidence indicates that the two verbs are not simply coordinated, but that they form a single syntactic unit. For example, as first pointed out in Ross (1986 [1967]) the construction allows a violation of what generative grammarians have called the 'island constraint' (or more recently, 'subjacency'). For example, it is possible to say *Who shall we go and see on Sunday*, where the wh-pronoun has been 'moved out' of the coordinate structure *go and see who*, indicating the unitary syntactic status of *go and see*.

Let us now turn to the semantics of the construction. Consider examples (1) through (4), which show some typical examples from English:[1]

(1)   a.   Look what you've gone and done!
        b.   He's gone and lost his job.
        c.   It was going to be a surprise, but he went and told her.
(2)   Nobody thought he could climb Everest, but he went and did it!
(3)   We asked him not to call the police, but he went (ahead) and did it anyway.
(4)   a.   I think we should all go and see Valerie on Sunday.
        b.   I'll go and get the rest of your stuff.

These examples show that the *go-and-Verb* construction occurs in a variety of uses: in examples (1a-c) it seems to express 'annoyance on the part of the speaker,' an implication that the action described by the main verb is 'stupid' or 'undesirable;' in example (2) it expresses a certain degree of 'surprise;' in example (3) it conveys something like 'proceeding without hesitation' or 'without regard to others;' in examples (4a-c) it expresses 'actual motion.'

Note also that in example (3) there is an optional particle, *ahead*. I will briefly return to the place of such particles in the overall account developed in this paper at the end of Section 4.

The wide range of uses associated with the *go-and-Verb* construction has prompted most authors who mention its meaning at all to react in one or both of the following ways: they either concentrate on the uses where *go* is used literally, or they treat the

---

[1] All examples in this paper are constructed. For English, I have based them on an analysis of some 200 examples from the *OED* and various corpora (I have chosen to use constructed examples rather than the original data for expository ease; cf. Stefanowitsch 2000 for an analysis of English corpus data). For the other languages, I work with what my consultants (and in some cases, the relevant literature) have provided.

different uses as unrelated idiomatic expressions, usually mentioning only one of the uses (for example, most dictionaries mention only examples like those in [1]), or claiming that *go* has lost most of its meaning in such expressions (e.g. Visser 1969).[2]

However, both approaches fail to account for a fact which has largely gone unnoticed, namely that motion verbs like *go* in general, and the *go-and-Verb* construction in particular, exhibit some cross-linguistic semantic regularities that are rather surprising if one assumes that *go* either means 'go' or nothing at all.

First, *go* and other basic motion verbs are used in many verb serializing languages in order to impose a motion profile onto an otherwise stative verb, or to give other motion verbs a deictic orientation, with *go* typically expressing motion through space in general or away from the speaker in particular (cf. Sebba 1987), as in the following example from Akan:

(5)   oguaɛ   kɔɔ   ahabaɛ   mu
        he-flee-Past   go-Past   bush   in
        'he fled into the bush' (Sebba 1987)

In such constructions, then, the motion verb adds an aspect of motion or deixis to the overall meaning of the expression. Note that this is exactly the function of *go* in examples 4a-b above: neither *see* nor *get* have an inherent motion component to their meaning. In both examples, *go* imparts a motion reading onto an otherwise stative verb (or at least one not associated with motion through space). Both situations can alternatively be described with verbs that *do* have such an inherent motion component: *go and see* means roughly the same as *visit*, and *go and get* means roughly the same as *fetch*.

Second, *go* and related verbs often serve as a source for tense and aspect morphology in grammaticization processes. Bybee et al. (1994) have shown that *go* develop into markers for progressive, continuative, or habitual aspect (as well as future tense, and sometimes perfective aspect); such aspectual notions are also encoded by the *go-and-V* construction across languages (see section 4 below, cf. also Stefanowitsch 2000).

Third, other languages have structurally parallel constructions which have some of the same semantics associated with them as the English examples in (1) through (4) above (again, by assuming that *go* contributes nothing to the *go-and-Verb* construction, this cross-linguistic pattern remains unaccounted for).

Consider the following examples from Swedish:

(6)   a.   *Han*   *har*   *gått*   *och*   *gift*   *sig.*
                he   has   gone   and   married Refl.
                'He went and got married.' (Joseffson 1991)

---

[2] The semantics of the *go-and-Verb construction* is mentioned or referred to in passing in Carden and Pesetzky (1979), Kruisinga and Erades (1947), Poutsma (1928), Shopen (1971), Visser (1969), Lakoff (1986), Quirk et al. (1991), Zandvoort (1966); for Swedish see Joseffson (1991), cf. also Hulden (1961), Wellander (1965); for Spanish see Arnaiz and Camacho (1999).

b.   *Och*   *så*   *går*   *han*   *och*   *berättar*   *det*
      and   so   goes   he   and   tells   it
      för   *sin*   *fru!*
      to   his   wife
      'And then he goes and tells it to his wife.' (Joseffson 1991)

c.   *något*   *jag*   *har*   *gått*   *och*   *tänkt*   *mycket*   *på*
      sth.   I   have   gone   and   thought   much   about
      'something I have been thinking about a lot'

The meanings commonly associated with this construction are 'surprise' and/or 'disapproval' in (6a-b), two meanings also found for the English construction. Example (6c) shows a use of the construction which is not found in English: 'activity for an extended period of time,' corresponding functionally to the progressive in English.

Structurally and semantically parallel examples can be found in Danish (unless otherwise indicated, all non-English examples are from the consultants mentioned in the acknowledgment note):

(7)   a.   *Lad*   *os*   *gå*   *hen*   *og*   *danse*   *i aften.*
           let   us   go   there   and   dance   tonight
           'Let's go and dance tonight.'

       b.   *Han*   *var*   *gået*   *hen*   *og*   *havde*   *gifted*
           He   had   gone   there   and   had   married
           sig.
           Refl
           'He went and got married.'

(8)   *Jeg*   *går*   *(rundt)*   *og*   *tænker*   *på*   *den*
     I   go   (around)   and   think   about   that
     eksamen   *hele*   *tiden.*
     exam   all   the-time
     'I am thinking about that exam all the time.'

Again, the construction is used to express 'actual motion' in (7a), 'surprise' and/or 'disapproval' in (7b), and 'continuous action' in (8). Notice that there is an obligatory particle *hen* 'there, over' in the construction expressing the first two meanings, and an optional particle *rundt* 'around' in the construction expressing the third meaning. I will return to the importance of these particles in Section 4.

Next, consider the following example from Spanish (Arnaiz and Camacho 1999):

(9)   *Y*   *entonces,*   *el*   *niño*   *va*   *y*   *se*   *cae.*
     and   then   the   boy   goes   and   Refl   falls
      'And then, the boy (suddenly/unexpectedly) falls.'

The meaning commonly associated with this construction is a 'sudden change in the expected flow of events' (e.g. in a narrative), i.e. 'unexpectedness' or 'surprise.'

Finally, consider these examples from Modern Hebrew:

(10) a.   *Kulam*   *paxadu*   *liftoax*   *et*
             everyone   was-afraid   to-open DO
             ha-kufsa,   *aval*   *Dan*   *halax*   *ve*   *asa*   *et*
             the-box   but   Dan   went   and   did   DO
             ze
             it
             'Everyone was afraid to open the box, but Dan just did it.'

      b.   *Dan*   *halax* *ve*   *kana* *lo*   *etmol*   *shaon*
             Dan   went   and   bought him   yesterday   watch
             xadash.
             new
             'Dan went and bought himself a new watch yesterday.'

The meanings commonly associated with this construction are the idea of a 'daring act', of 'no attention paid to obstacles,' as in example (10a); in addition, it can express 'surprise,' i.e. (10b) would be used in a situation where the act of buying a watch is for some reason unexpected.

In sum, *go* is being put to many different uses across languages.[3] However, these uses repeat themselves across languages: *go*, and more specifically the *go-and-Verb* construction, have certain cross-linguistically recurring uses, which include expressing:

- actual motion through space;
- annoyance, disappointment, disapproval;
- evaluation of an action as stupid or unfortunate;
- surprise, sudden change in the expected flow of a narrative;
- proceeding without hesitation or regard to others, paying no attention to obstacles;
- continuous action, progressive/habitual aspect.

There are two conclusions to be drawn from this cross-linguistic recurrence of certain meanings: first, and most importantly, *go* is not being used in arbitrary ways. Instead, there seems to be an underlying systematicity to the semantics associated across languages with constructions containing a general motion verb in conjunction with another verb. Second, this systematicity must in some way be related to the meaning of the verb *go*.

We are confronted with the issue of how to characterize the semantics of such general motion verbs in a way that allows insights into the current problem. After all, we are dealing with different words from different languages. Specifically, we have to find a

---

[3] The survey given here is far from complete. Constructions which are structurally parallel or closely parallel to the *go-and-Verb* construction can be found, for example, in Portuguese, Greek, Finnish, and Japanese. A thorough investigation of the functions of motion verbs in verb-serializing languages also seems promising: for example, in Supyire, *sa* 'go' can mark the beginning of a new action after some action which has gone on for some time, cf. Carlson (1994: 305).

way of capturing the semantics of verbs like *go* language-independently. I believe that this can be done in terms of *image schemas* (in the sense of Johnson 1987).

3. THE IMAGE-SCHEMATIC STRUCTURE OF MOTION VERBS.

Image schemas are defined as general patterns which we abstract over recurrent experiences, they 'emerge as meaningful structures for us chiefly at the level of our bodily movements through space, our manipulation of objects, and our perceptual interactions' (Johnson 1987: 29). Such schemas are non-linguistic, involve all sensory modalities, and there is evidence that they are acquired during the first year of life, i.e. before language acquisition proper sets in (cf. Mandler 1992).

An example for such a schema is given in Figure 1: the *motion schema*. In its most basic form the motion schema specifies a trajector moving along a path which functions as its landmark. A child acquires this schema by extracting commonalities over many instances of the experience of moving (or being moved) through space and of watching things move through space.
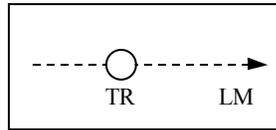

FIGURE 1. THE MOTION SCHEMA

Their non-linguistic nature as well as the fact that they are grounded in basic experience not specific to a particular culture make image schemas ideal candidates for cross-linguistic semantic descriptions: they are evoked as a central aspect of the meaning of any given word. Note that this does not mean that image schemas can be equated with word meaning; the semantics of any given word will include much more than just image schematic structure. For example, the English verbs *crawl* and *run* both evoke the motion schema, but in addition they invoke, for example, information about speed ('slow' vs. 'fast').

Thus, we can assume that the English verb *go* and its nearest equivalents in other languages—although they may differ in their specific semantics—evoke as part of their meaning the motion schema in Figure 1.

Since motion events may differ from each other in certain fundamental ways, there is a number of more richly specified variants of the basic motion schema which are likely to be shared across cultures. These variants can be seen as schemas of intermediate abstractness, including more information than just the configuration of a trajector moving along a path. Figures 2a-c show what I take to be typical variations of the motion schema.
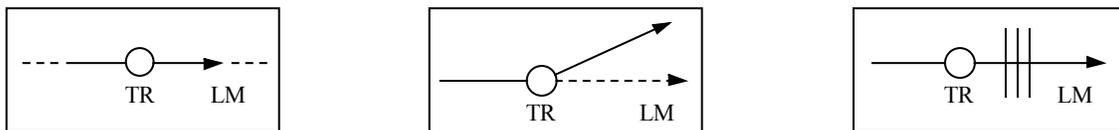
(a) *Extended motion*   (b) *Divergence*   (c) *Potential obstacles*

FIGURE 2. VARIANTS OF THE MOTION SCHEMA

Figure 2a contains additional information about duration: it is a schema of motion over an extended period of time (symbolized by the broken lines on either side of the arrow. Figure 2b shows a situation where the actual path of the trajector (symbolized by the solid arrow) diverges from the expected path (symbolized by the broken arrow). Figure 2c shows a variant of the basic schema which focuses on potential obstacles (symbolized by the vertical lines), which have to be overcome by the trajector.

Image schemas do not just organize basic bodily experience; they can also be extended to structure abstract thinking via conceptual metaphors. In the case of the *go-and-Verb* construction, the basis for extension is the metaphor PROCESSES (or ACTIONS) ARE MOTION, a submetaphor of the very widespread metaphor CHANGE IS MOTION.[4]

The literal and the extended uses of the basic schema in Figure 1 and its variants in Figure 2 can account for the semantic contribution of *go* to the *go-and-Verb* constructions:

- 'actual motion' is the basic meaning of all schemas in Figures 1 and 2;
- meanings like 'continuous action' in the Scandinavian languages and progressive/habitual aspect more generally can be accounted for in terms of a metaphorical extension of the schema in Figure 2a from 'motion over an extended period of time' to 'action over an extended period of time';
- meanings like 'surprise' and 'unexpectedness' result from the same metaphorical extension applied to the schema in Figure 2b. Here, the expected conceptual path corresponds to the expected flow of events (in real life or in a narrative), and the divergence from this path corresponds to an unexpected event;
- meanings like 'annoyance,' 'disappointment,' 'disapproval,' or 'evaluation of an action as stupid' are also accounted for by the schema in Figure 2b, with the difference that the expected conceptual path corresponds to the kinds of actions that the speaker considers desirable or rational, and the divergence from this path corresponds to an event that is not desirable or rational from the speaker's perspective;
- finally, meanings like 'proceeding without hesitation,' 'paying no attention to obstacles' can be accounted for in terms of the same metaphorical extension as before, this time applied to the schema in Figure 2c, where the obstacles correspond to potential reasons not to act in a certain way.

4. IMAGE-SCHEMA BLENDING.

---

[44] This metaphor is described in detail in Radden (1996), it is entailed by the metaphor STATES ARE LOCATIONS, which is exemplified by expressions like *John is IN love*. A simple example for the metaphor CHANGE IS MOTION is *John WENT crazy*, an example for ACTION IS MOTION is *He STOPPED SHORT OF hitting her*, or *I don't agree with this COURSE of action*.

So far I have shown (part of) the image-schematic structure evoked by *go* and its equivalents in other languages, and I have shown how this structure can account for the semantics of the constructions presented in Section 2. I have not yet addressed the issue of how the *go-and-Verb* construction as a whole works, i.e. how the semantics of *go* are combined with that of the main verb. Section 4.1 deals with this issue; Section 4.2 points out further applications of the mechanism proposed.

4.1. The *go-and-Verb* Construction as Image-schema Blending. I propose that the *go-and-Verb* construction fuses the semantics of *go* and the main verb into a single event frame. The *and* in *go and Verb* does not function as a coordinator, but as a semantic instruction to blend (a variant of) the image-schematic structure evoked by *go* with the event structure evoked by the main verb. This integration allows the speaker to construe the event in accordance with the motion schema.

Consider example (2), repeated here as (11):

(11)   Nobody believed he could climb Everest, but he went and did it.

The event encoded by *he climbs Everest* is an instance of the transitive event schema, in which an agent acts on a patient with some result. By blending this schema (or rather, one of its semantically richer manifestations) with the divergence schema (shown in Figure 2b above), the event is construed as a divergence from an expected conceptual path—in this case, an expected course of action which would have involved the agent doing nothing to the patient (i.e. *not climbing Everest*). This analysis is summarized in Figure 3.
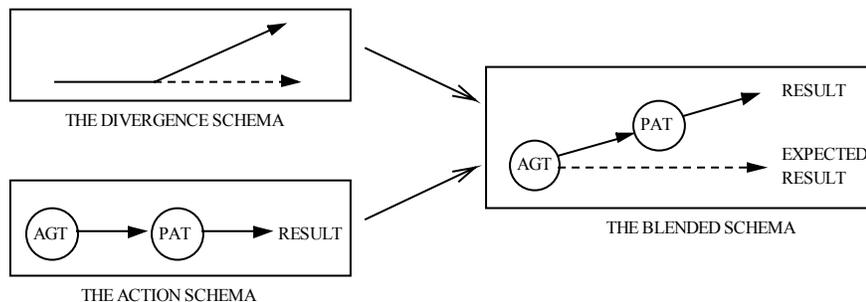

FIGURE 3. BLENDING THE DIVERGENCE SCHEMA AND THE (TRANSITIVE) ACTION SCHEMA

It is not clear how much this sense of *blending* shares with that developed by Gilles Fauconnier and Mark Turner (cf. Fauconnier 1998 for an overview). In their sense of blending, the blend (more precisely, the blended space) contains a selected subset of elements from each of the sources (the input spaces), as well as additional elements not present in either of the input spaces. In contrast, the sense of the term used here refers to a complete incorporation of the image-schematic structure of *go* into the more richly specified event frame of the main verb, which is why I also refer to it as *fusion* (cf. Stefanowitsch 2001 for further discussion of this issue).

Evidence for such a complete incorporation of the two schemas comes from the syntactic properties of the construction. Recall that the two conjoined verbs behave syntactically as a single unit. This makes sense: if semantically they *are* a single unit, we would expect this unity to be reflected on the formal side.

The particles which appear in some of the examples can be naturally integrated into this account: their function is to specify a particular variant of the motion schema explicitly. English uses *ahead* to pick out the schema in Figure 3c, thus it occurs in examples that express the fact that someone does something with no regard to potential obstacles (for a more detailed discussion of this and other particles occurring in the English *go-and-Verb* construction, cf. Stefanowitsch 2000). Danish uses *hen* 'there, over' to pick out the schema in Figure 2b, i.e. to express deviation from a conceptual path (i.e. an expected or desirable course of action). Danish can also use *rundt* 'around' to specify the variant of the motion schema that specifies motion for an extended period of time. Intuitively, the particular particles used seem to be well motivated by the corresponding schemas.

Finally, the image-schema blending account is compatible with what is known about the discourse function of the *go-and-Verb* construction. Sanchez (1999) has shown that the *go-and-Verb* construction and the *come-and-Verb* construction are typically found in three discourse contexts: (i) commands, suggestions, and invitations; (ii) introducing a new episode in a narrative; and (iii) returning to the main topic in a narrative after a digression. All of these functions can be related to the divergence schema shown in Figure 2b above: commands, suggestions, and invitations all entail a divergence from the expected flow of events (in Speech Act Theory, it is actually one of the felicity conditions of such speech acts that the hearer was not already planning to do what the speaker suggests he do, cf. Levinson 1983: 240), and a digression or a return to the main topic within a narrative are also divergences from the currently expected flow of events.

4.2. Some Other Applications of Image-schema Blending. The account given in the preceding section has the potential of being applied to many other areas of language than that for which it was developed here. I will mention three areas here, but of course, this is not an exhaustive list.

First, the account can be extended to other *Verb-and-Verb* constructions which can be found in English and other languages, with main verbs such as *sit*, *stand*, *run*, and *try*. For example, let me briefly comment on the *try-and-Verb* construction. According to Quirk et al. (1991: 978), *I'll try and come tomorrow* is 'roughly equivalent to *I'll try to come tomorrow*, but is more informal in style.' However, as Nordquist (1998) has shown, this is not true: *try and* is more likely to be used when the action described by the main verb is unlikely to be completed, whereas *try to* is neutral with respect to this parameter. This makes sense in terms of image-schema blending: the *try-and-Verb* construction blends the semantics of both verbs into a single event schema. The event structure of this event will incorporate the inherent semantics of 'incompletion' provided by *try and* into the event structure of the main verb. The *try-to-Verb* construction on the other hand does not blend the semantics of the two verbs; it simply encodes two separate events, one of which is the purpose of the other. Thus the event encoded by the main verb does not incorporate the potential incompleteness of *try*.

Second, a blending account also proves insightful for analyzing the English *Adjective-and-Adjective* construction (as in *I'll come when I'm good and ready*, *It's nice and warm in* here, etc.). Young (1999) and Young and Stefanowitsch (2000) have drawn on the mechanism postulated in the preceding section to show that such constructions do not encode two separate qualities (as the co-ordinating conjunction suggests), but that they encode a single quality which is a blend of the two conjuncts. Thus, for example, *It's nice and warm in here* does not mean 'it is both nice and warm in here,' but it means 'it is warm in here and the warmth is nice.'

Finally, the idea of image-schema blending seems to have great potential in the analysis of various aspects of sign language. Sign languages are less restricted than spoken languages in terms of linearization; if image-schema blending is indeed a universal conceptual mechanism, it seems that sign languages are structurally very well suited for exploiting it.

An example may clarify this idea and show that it may indeed be right: in American Sign Language, a whole array of aspectual distinctions is marked by imposing different kinds of motion schemas onto the same basic hand shape encoding a verb or an adjectival predicate (cf. Klima and Bellugi 1979). For example, durational aspect is encoded by 'smooth, circular, reduplicated movement,' and continuative aspect is encoded by 'slow, elongated, continuous reduplications that are elliptical in shape' (ibid.: 294). In other words, ASL does exactly the same thing that Danish is doing when it blends a verb with *gå rundt* 'go around' (as in [9] above). The difference is that ASL can create a perfect blend due to the possibility of simultaneity, while Danish has to indicate this simultaneity with the conjunction *og* 'and'.

5. CONCLUSION.

This paper has offered a unified account of the *go-and-Verb* construction and its meaning across languages. It was argued that all uses of this construction are motivated by the image schematic properties of the verb *go* and whichever second verb occurs in a particular expression. More specifically, the image-schematic properties of *go* are incorporated into (or blended with) the event structure of the second verb to allow the speaker to construe the event denoted by the second verb in accordance with the image-schematic meaning of *go*.

The account of the *go-and-Verb* construction in terms of image-schema blending suggests a motivation for the fact that these structurally and semantically similar constructions exist in different languages.

The semantic similarity of these constructions is due to the fact that the conceptual structure proposed here is independent of language; the slight variation in the exact set of meanings associated with it in any particular language is due to the fact that conceptual structure manifests itself in language-specific ways.

The fact that the constructions are formally similar is due to the fact that every element in the construction is motivated by some aspect of conceptual structure, with the main verb encoding some event structure, *go* encoding some construal in accordance with its image schematic meaning, and a coordinating conjunction encoding the blending of

the two (i.e. the incorporation of the image-schematic structure of *go* into the event structure of the main verb).

REFERENCES

ARNAIZ, ALFREDO, and CAMACHO, JOSE. 1999. A Topic Auxiliary in Spanish. *Advances in Hispanic Linguistics*, ed. by F. Martinez-Gil and J. Gutierrez-Reixach. Somerville, MA: Cascadilla Press.

BYBEE, JOAN, REVERE PERKINS, and WILLIAM PAGLIUCA. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago and London: The University of Chicago Press.

CARDEN, GUY, and DAVID PESETZKY. 1979. Double-verb Constructions, Markedness, and a Fake Coordination. *Chicago Linguistic Society* 13.82-92.

CARLSON, ROBERT. 1994. *A Grammar of Supyire.* Berlin and New York: Mouton de Gruyter.

CORMACK, ANNABEL and NEIL SMITH. 1994. Serial Verbs. *UCL Working Papers in Linguistics 6*, University College London.

EKBERG, LENA. 1993. The Cognitive Basis of the Meaning and Function of Cross-linguistic Take and V. *Perspectives on language and Conceptualization. (Belgian Journal of Linguistics 8),* ed. by Jan Nuyts and Eric Pederson, 21-42.

FAUCONNIER, GILLES. 1998. Mental Spaces, Language Modalities, and Conceptual Integration. *The New Psychology of Language: Cognitive and Functional Approaches to Language*, ed. by Michael Tomasello, 251-279. Mahwah, NJ and London: Lawrence Erlbaum.

HULDEN, LARS. 1961. Om samordning hos Bellman. *Studier i Nordisk Filologi* 51.7-149.

JAEGGLI, OSVALDO, and N. HYAMS. 1993. On the Independence and Interdependence of Syntactic and Morphological Properties: English Aspectual COME and GO. *Natural Language and Linguistic Theory* 11(2).313-346.

JOHNSON, MARK. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: The University of Chicago Press.

JOSEFFSON, GUNLOG. 1991. Pseudocoordination - a VP + VP Coordination. *Working Papers in Scandinavian Syntax 47*, 130-156. Lund: University of Lund, Dept. of Scandinavian Languages.

KLIMA, EDWARD S. and URSULA BELLUGI. 1979. *The Signs of Language*. Cambridge, MA and London: Harvard University Press.

KRUISINGA, ETSKO, and P.A. ERADES. 1947. *An English Grammar. Vol. I, Part 1: Accidence and Syntax.* Groningen: Noordhoff.

LAKOFF, GEORGE. 1986. Frame Semantic Control of the Coordinate Structure Constraint. *CLS* 22(2).152-167.

LEVINSON, STEPHEN. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

MANDLER, JEAN M. 1992. How to Build a Baby: II: Conceptual Primitives. *Psychological Review* 99.587-604.

NORDQUIST, DAWN. 1998. Try and: A Discourse Analysis. *Proceedings of the First High Desert Linguistics Society Conference*, April 3-4, 1998, Albuquerque, NM.

POUTSMA, HENDRIK. 1928. *A Grammar of Late Modern English. Part I, Second half: The Composite Sentence*. 2nd edition. Groningen: Noordhoff.

PULLUM, GEOFFREY K. 1990. Constraints on Intransitive Quasi-serial Verb Constructions in Modern Colloquial English. *When Verbs Collide: Papers from the Ohio State Mini-conference on Serial Verbs.* (Ohio State University Working Papers in Linguistics), ed. by Brian D. Joseph, Arnold M. Zwicky, 218-239.

QUIRK, RANDOLPH, SIDNEY GREENBAUM, GEOFFREY LEECH, and JAN SVARTVIK. 1991. *A Comprehensive Grammar of the English Language.* Ninth, revised impression. London and New York: Longman.

RADDEN, GÜNTER. 1996. Motion Metaphorized: The Case of Coming and Going. *Cognitive Linguistics in the Redwoods. The Expansion of a New Paradigm in Linguistics*, ed. by Eugene H. Casad, 423-458. Berlin and New York: Mouton de Gruyter.

ROSS, JOHN ROBERT. 1986. *Infinite Syntax*! Norwood, NJ: Ablex [= 1967. *Constraints on Variables in Syntax*, Ph.D. Thesis, MIT].

SANCHEZ, IVO. 1999. English Pseudo-coordination. Paper presented at the Second Annual High Desert Student Conference in Linguistics, Albuquerque, NM: University of New Mexico, March 26-28, 1999.

SEBBA, MARK. 1987. *The Syntax of Serial Verbs*. Amsterdam: Benjamins.

SHOPEN, TIMOTHY. 1971. Caught in the Act. *Chicago Linguistic Society* 7.254-263.

STAHLKE, HERBERT. 1970. Serial Verbs. *Studies in African Linguistics* 1.60-99.

STEFANOWITSCH, ANATOL. 2000. The English *go-(PRT)-and-VERB* Construction. *Proceedings of the Twenty-sixth Annual Meeting of the Berkeley Linguistic Society*, February 18-21, 2000, University of California, Berkeley.

STEFANOWITSCH, ANATOL. 2001. Redemptive Aspect in English: The *go-back-and-VERB* Construction. Paper presented at the 7th International Cognitive Linguistics Conference, Santa Barbara, CA: UC Santa Barbara, July 22-27, 2001.

VISSER, F. TH. 1969. *An Historical Syntax of the English Language. Part III.First Half: Syntactical Units with Two Verbs*. Leiden: Brill.

WELLANDER, ERIK LUDVIG. 1965. *Riktig svenska: en handledning i svenska sprakets vard.* Third edition. Stockholm: Norstedt.

WIKLUND, ANNA-LENA. 1996. Pseudocoordination is Subordination. *Working Papers in Scandinavian Syntax 58*, 29-53. Lund: University of Lund, Dept. of Scandinavian Languages.

YOUNG, HILARY. 1999. The Adj+Adj construction in English. Manuscript. Houston, TX: Rice University.

YOUNG, HILARY and ANATOL STEFANOWITSCH. 2000. Domain Blending in English: the *ADJ-and-ADJ* Construction. Paper presented at the Conference on Conceptual Structure, Discourse, and Language, Santa Barbara, CA: UC Santa Barbara May 11-14, 2000.

ZANDVOORT, REINARD W. 1966. *A Handbook of English Grammar*. 3rd edition. Englewood Cliffs, NJ: Prentice-Hall.

FOCUS AND QUOTATIVE *LIKE*: A NECESSARY DICHOTOMY?

ANDREW TISTADT

University of New Mexico

1. INTRODUCTION.

   The discourse marker *like* has received little attention during its relatively brief presence in the English language. Schiffrin does not mention it in her 1987 work *Discourse Markers* and a quick perusal of linguistic or MLA indices turns up only a handful of relevant articles. Nevertheless, nonstandard *like* proves a fascinating topic of study, and, although not all researchers agree on its polysemous nature, it is generally divided into two subcategories—focus marker and quotative complementizer:

   (1) She cut her hair *like* real short. [focus marker]
   (2) And she was *like*, "I just can't believe it." [quotative complementizer]

Blyth, Recktenwald, and Wang (1990) offer the term FOCUS QUOTATIVE as a possible alternative label for *quotative complementizer* since in their opinion, "it is entirely plausible to interpret [quotative] *be like* as a variant of the focus *like*" (224). (The term DIALOGUE INTRODUCER has also been used by scholars [Ferrara and Bell 1995].) Focus marker *like* has created less controversy, which is perhaps why most articles regarding nonstandard *like* deal mainly with *like* as dialogue introducer (Ferrara and Bell 1995, Blyth et al. 1990, Romaine and Lange 1991, Meehan 1991). In fact, to my knowledge, the only study dedicated to focus *like* so far is Underhill (1988). Of course, none of the other scholars completely ignores the focus function of *like*, but generally the studies have been one-sided. In the following paper, then, I intend to examine and discuss the many uses of both functions of *like*, as well as propose that Blyth et al. (1990) are correct in viewing *be + like* as a simple variant of focus *like*.

   Before I begin my analysis of nonstandard *like*, I would first like to discuss the concept of discourse marker. Schiffrin (1987) defines it as a "sequentially dependent [element] which brackets [a] unit of talk" (31). Her examples include the following items: *oh*, *well*, *and*, *but*, or, so, *because*, now, then, *I mean*, and *y'know*. Her definition, however, does not seem as clear or precise as that offered by Romaine and Lange (1991), which states: "discourse markers [are those] particles which are used to focus on or organize discourse structure" (245). Focus *like* may be considered a discourse marker, then, because it "has no semantic content of its own" (Underhill 1988:245) and simply serves to aid the speaker in discourse organization by allowing him or her to highlight the most important part of his or her utterance. Likewise, because focus *like* is semantically empty, it can generally be removed from an utterance without much affecting its grammaticality or meaning (see [1] above). This in turn supports its discourse marker status, for one of Schiffrin's (1987) properties of discourse markers is that they "are

never obligatory" (64). Of course, this property creates problems for labeling *be + like* as a discourse marker since removing *like* in cases such as (2) would render the sentence ungrammatical. Romaine and Lange (1991) recognize this discrepancy and consequently contend that the entire "quotation frame" (246) would need to be deleted in order for the aforementioned property of markers to be met (although by doing so, I would argue, the meaning of the utterance would indeed be modified since the context for the quote would be lost). Thus in an utterance such as (2), the entire phrase *be like* must be considered the discourse marker, as opposed to focus marker *like*, which stands on its own.

The data for this study were gathered by recording two Anglo females in their early twenties as they drove around the streets of Albuquerque, New Mexico. The topics ranged from work to guys, from plans for the evening to discussion of friends. Both females are college-educated, members of the urban middle class, and apparently are good friends as well as co-workers. The conversation seemed fluid, natural, and unforced, and the knowledge that their utterances were being captured for a linguist's perusal did not seem to affect their performance. All examples below, unless otherwise specified, are from these data.

2. QUOTATIVE *LIKE*.

As previously stated, most research involving nonstandard *like* has concentrated on its role as dialogue introducer. When introducing dialogue, many options are at the speaker's disposal. Most common is the mundane *say*, along with more descriptive verbs of saying such as *reply*, *respond*, *shout*, *whisper*, *exclaim*, *growl*, and many others generally only employed by fiction writers. Often, however, these types of *verba dicendi* seem too formal or literary for everyday speech, so many speakers choose to introduce dialogue with one of the following nonstandard quotatives: *go*, *be + all*, *be + like*, or even just *be*. Example (3) provides an example of historical present *go*.

> (3) And then I *go*, "What d'ya mean you don't like syntax?" And she
>   *goes*, "I don't know. It's just not my thing." [example mine]

The fictitious speaker in (3) could have chosen *be + like* or *be + all* rather than *go* in order to introduce the dialogue, but not because *go* and *be + like/all* are completely synonymous. Indeed, unlike *go*, *be + like* and *be + all* are capable of introducing gestures, as well as quotation, as the following example from the data indicates:

> (4) ...whenever I come around he's *like*...[non-verbal gesture]

Since my data were captured on audiotape, I am unaware of what exact gesture the speaker in (4) made, but because of certain audible clues, I do know that she made some sort of gesture. The point, however, is that in (4), *go* could not be used in the same manner as *be + like* (although *be + all* seems a grammatical option).

Another difference between *go* and *be + like/all* is clearly seen when the latter is used to introduce internal monologue, i.e., thought, as we see in (5):

> (5) Susan like jumped up and ran out there. She's *like* "Who's got the beer?" [laughter]. I was *like* "Go Susan". [reporting an incident at work when a naive client comes in office with beer in hand]

Because of the context in which this reported dialogue takes place (a formal office), it is not likely that the speaker in (5) actually uttered the words *Go Susan*. Rather, the speaker is reporting a thought which occurred to her at the time Susan ran out and confronted the beer-drinking client. If the speaker had uttered, "And I *went*, 'Go Susan'," we would have no choice but to assume that she did in fact say these words because of *go*'s inability to introduce internal monologue. Notice, however, that "I was *all*, 'Go Susan'" also seems capable of introducing thought, but perhaps to a lesser degree than *like*. Additionally, both *like* and *all* are at times equally ambiguous in whether they refer to internal or external dialogue: only context, as in this case, permits correct interpretation. More research is needed to determine what semantic differences, if any, exist between *be + all* and *be + like*.

The following excerpt from the data provides many examples of quotative *like*:

> (6) But, she's *like*, "Will you come be my maid-of-honor, or witness, or whatever the you want to call it." And I was *like*, "OK." And then he called and he's *like*, he goes, "Well, uh, I-I need to pick out a ring for Yvette, and I figured since y'all had been friends for like, uh, like a long time, like 10 years, that, you'd kinda know what she'd like, and would be able to help me out 'cause..." This is exactly how he sounds on the phone, like a big-ass drawn-out message. He's *like*, "...and, and I think she'd like it because what better friend than her best friend to help me pick out for her." He's *like*, "Together we'll pick out something she likes." And I'm *like*, I'm *like* sitting there listening I'm *like*, uh-huh, uh-huh...

Obviously, not all instances of *like* in this passage are quotative, but, for the moment, we will only focus on those that are (instances of lexical and focus *like* will be ignored). First of all, we notice that the speaker tells this anecdote almost completely in the historical present, a common device used to enliven discourse and one to which *be + like* readily lends itself since all forms of present tense *be* permit contraction (*I'm*, *you're*, *he's*, *she's*, *it's*, *we're, they're*). Second, we see that the speaker freely switches between first and third person while using *like* to introduce direct speech, thereby confirming other researchers' claims of this possibility (*Cf.* Ferrara and Bell 1995). Also worth noting is that in line 3 we see apparent evidence for *go* and *like*'s occasional synonymity. I would argue, however, that because of *be + like*'s previously explained ambiguity regarding internal versus external speech (and *go*'s lack of it), these two quotatives, even in the context of line 3, are not completely synonymous. The use of *go* after *like* allows the speaker to emphasize the external nature of the following reported dialogue, whereas *like*

by itself would leave the following reported speech ambiguous. Consider, for example, the semantic difference between (7) and (8):

(7) And she *goes*, "Who cares?"
(8) And she was *like*, "Who cares?" [examples mine]

There should be no question in the reader's mind that (7) does not allow for the possibility of internal monologue. Example (8), on the other hand, seems completely ambiguous due to lack of context. It could either mean, "And she was obviously thinking, 'Who cares?'," or "And she literally said, 'Who cares?'"

The reason for *like*'s ambiguity and *go*'s lack of it may be explained, I believe, by addressing their original meanings. *Go*, being a verb of movement, originally meant "to move or travel," and according to *The American Heritage College Dictionary* (1993:583), "has long been used to describe the production of nonlinguistic noises, as in *The cow goes 'moo.'* In recent years, however, younger speakers have extended this use of *go* to the report of speech, as in *Then he goes, 'You think you're real smart, don't you?'* For speakers young enough to get away with it, this usage serves a useful purpose as an explicit indicator of a direct quotation [and] is highly inappropriate in formal speech or writing." I am not sure young speakers are the only ones who "get away with" using *go* in this manner, for *go* as a quotative has been around much longer then *be + like* (see Butters 1980 and Schourup 1982 for early research on quotative *go*). In fact, in their quantitative study of *be + like*, Blyth et al. (1990) did not find anyone over the age of 38 (out of a range of 20-72 years) using quotative *like*, thus implying that if the more recent *be + like* is being used (in 1990) by 38-year-olds, it is likely that the older quotative *go* has spread even further through the generations.

So what lies behind quotative *like*'s ambiguity? Tannen (1986:321) offers a promising, albeit somewhat contradictory explanation when she writes "[*Be + like* is] a formulaic introducer, not by its literal meaning, but simply by convention. If the literal meaning functions at all, it is to suggest that the dialogue is not being quoted but simply represents the kind of thing that character was saying or thinking." I would argue, however, that the literal meaning of *like* ('similar to') is always present when used as a dialogue introducer, meaning that the speaker continually has the option of quoting direct speech or internal monologue. This would first explain *like*'s occasional ambiguity in this regard and, second, accurately reflect the oft-proposed idea that reported dialogue is merely an approximation of what was actually said. ("The variable [*be + like*]...shows sensitivity to the well-known fact that quotations are representations rather than verbatim reproductions" [Ferrara and Bell 1995:286]. Tannen's term for quotations, CONSTRUCTED DIALOGUE [1986], likewise reflects this idea.) Of course, Tannen is correct in asserting that *be + like* has acquired a certain formulaic feel to it, as (6) demonstrates, but because of the frequency with which the variable is also used to introduce internal monologue, I would argue that it still retains much of its original meaning. Consider, for example, (9):

(9) He *like* looked *like* he was *like* "Oh fuck, I shouldn't uh answered the
    door. I feel like such an idiot." [referring to an instance when speaker
    pays a surprise visit to an improperly dressed male at his home]

Amazingly, the speaker in (9) uses *like* in three different ways. The first instance of *like* falls under the category of focus marker (the function of *like* we have yet to discuss, but essentially the absence of the verb *be* in this instance indicates that what immediately follows it will be important information rather than dialogue). Second, the *like* following *look* belongs to the lexical use of 'similar to.' The speaker wishes to communicate that the subject of her anecdote did not actually utter the following quotation, but appeared as if he could have been thinking it. In fact, the literal meaning of *like* here is what provides the proof that the third instance of *like* is internal monologue, rather than external. Had the surprised male actually said these words, the speaker would not have expressed the VP *looked like*, for this choice of words clearly indicates probability rather than verbatim reporting. Incidentally, the use of quotative *like* to introduce third person internal monologue was considered rare as recently as 1990 (Blyth et al. 1990), and even Ferrara and Bell (1995:272) pose the question (but provide no answer), "Do speakers report the inner thoughts of third persons with the variable [*be + like*]?" Example (9) answers this question affirmatively. Lastly, it is also worth noting that the speaker in (9) uttered these three cases of *like* without pause or hesitation, i.e., she did not say, "He like...looked like...," as if unsure of what to say next. Thus I do not think that the first or second uses of *like* in (9) can be classified as fillers or hesitaters. Rather, each plays a unique semantic role within the sentence.

Finally, there is one case of *be + like* which deserves special mention due to its ability to introduce hypothetical dialogue, viz. *it's + like*:

(10) It's *like*, "If you didn't want to have a beer, then, don't be wasting
    my 25 cents calling you."

The speaker in (10) is not reporting something that anyone actually said or thought, but rather, what one can now say/think in retrospect upon telling the story. As she narrates, the speaker adds her own perspective on what happened and, by using pleonastic *it* in subject position, she not only distances herself from the quotation, she communicates that it was not actually said or thought by anyone involved in the story. *It*'s *like*, then, allows the speaker to subtly add her own perspective on the reported events without specifically attributing that opinion to herself.

3. FOCUS *LIKE*.

Looking back at example (9), we see that the first instance of *like* lacks the verb *be*, which is indicative of *like*'s other nonstandard function as focus marker. Underhill (1988) defines focus as "the most significant new information in a sentence—often, the point of the sentence." Although at first glance focus *like* seems capable of appearing almost anywhere in a sentence, a closer look reveals a syntactic constraint—it is only used to introduce phrasal constituents (Underhill 1988). But of course, this gives focus *like* quite a bit of freedom since generally speakers' utterances are composed of many constituents. The following data provide clear examples of this (only focus marker *like* is in italics):

(11) You know she cut her hair *like* real short, ***like*** shorter than Susan's. [pre-AdvP, AdjP]

(12) Yeah, you know what I mean, *like* he's like a boy, *like* a boy, you know what I mean, ***like***, acts like a... [pre-IP, NP, VP]

(13) *Like* when I've gone over there and *like* peeked in those apartments...[pre-CP, VP]

(14) I don't like him either. He's like and he doesn't like me have you noticed *like* whenever he's like whenever I come around he's like...[pre-CP]

(15) You know, um...speaking of last names, um, when Duran was trying to pick her last name, um, she was going to pick Baca, which means 'cow,' *like* for a last name [pre-PrepP]

(16) Why don't we just ask Eric?...No, because then that'll look bad because he'll be like, "Who was askin' about you?"...I know he *like* totally has the hots for me. [pre-AdvP]

(17) He looked embarrassed, and *like* most guys wouldn't give a fuck, you know what I mean? [pre-IP]

The syntactic freedom of focus *like* arises from the speaker's ability to mark focus in a pragmatically flexible manner, that is, virtually any part of a sentence may be deemed as containing the most significant new information of that utterance. Look at (11), for example. The speaker here does not wish to merely point out that the antecedent of *she* has cut her hair, for that is an activity all humans participate in and does not merit our attention. Rather, the emphasis is placed on the fact that the woman changed hairstyles, from long to short, and that, for the speaker, this is the most important information of her utterance. The insertion of *like* prior to the AdvP *real short*, then, accomplishes the speaker's goal of highlighting this surprising information. (The second instance of *like*

here merely complements the main focus of the sentence; the comparison with Susan's hair allows the speaker to further specify the severity of the change.)

Example (16) also clearly demonstrates *like*'s ability to mark focus. The speaker wishes to emphasize the fact that Eric finds her attractive, rather than that she is aware of this. Thus, she places the focus marker after the initial IP, "I know...," but in front of the adverb *totally*, which, in turn, allows the speaker to stress the degree to which Eric feels attracted to her. All other examples of focus *like* in (11)-(17) demonstrate much the same emphatic usage.

I previously mentioned that *like*, when unaccompanied by *be*, indicates a focus function. However, sometimes *be* can appear alongside focus *like* without changing its function to that of dialogue introducer.

> (18) And he came to the door and he was *like* in his shorts
> (19) He just says like crusty things and he's *like* dirty
> (20) You know what I mean? Like, he's *like* a boy
> (21) I don't like the other dude that came in there. He's *like* controlling

The proof that these data exemplify focus and not quotative *like* can best be seen in (18). Were this an instance of quotative *like*, *like* would not be able to precede *was*, but as (22) shows, this is indeed possible:

> (22) And he came to the door and he **like** was in his shorts [example
>   mine] (Cf. *And he *like* was, "I don't understand." [example
>   mine])

The same is true for (19)-(21), although the placement of *like* before *is* would sound more marked due to present tense *be*'s tendency to contract when used with *like*. (In fact, I do not think present tense *be* + *like* is ever uttered without contraction, i.e., *I am like*, *He is like*, etc., which might indicate a process of grammaticization, reflected by the phonetic reduction in the construction.)

But are Blyth et al. correct in labeling quotative *like* a FOCUS QUOTATIVE, i.e.,, a simple variant of focus *like*? The data in (18)-(21) favor this idea, for they show how a transition between *be* + *like* as focus marker and *be* + *like* as quotative would not be difficult. Likewise, whenever a speaker introduces reported dialogue in narration, generally that quotation is the most important part of the utterance, i.e., the focus. Thus I agree with Blyth et al. in their assertion that quotative *like* may only be focus *like* in disguise. (It is interesting to note that the *American Heritage College Dictionary* [1993] does not mention a quotative function in its entry on *like*. They do say it can be used "to provide emphasis or a pause" as in *Like let's get going* (786), but this falls under the nonstandard category, as expected. Their reasons for ignoring quotative *like* are unknown: Is it too colloquial? too infrequent? Or perhaps just an inferable variant of focus *like?*)

We now might ask why *like* lends itself to use as a marker of focus. Romaine and Lange (1991) argue that "the meanings of 'approximative' and 'similarity' [of lexical

like]...have contributed to both discourse uses of *like*, and that these are natural outgrowths of existing uses and functions of *like*"(245). This supports Schiffrin (1987) when she points out that "Many discourse markers are used in ways which reflect their [earlier] meanings" (317). Thus a possible evolution of focus *like* could have begun with sentences such as, "Our team scored *like* 80 points" meaning they scored approximately 80 points, to more figurative uses of approximation, such as "That film was *like* sort of interesting" meaning "That film seemed interesting (but it wasn't really)." Next, *like* evolves to serves as a simple focus marker and thus acquires its current syntactic liberty ("That *like* was really interesting" meaning it indeed captured the speaker's attention). Finally, certain cases of focus marker *be + like* (see [18]-[21]) begin to be used in narrative to introduce dialogue, thereby giving rise to its current quotative function. A possible development, then, is as follows: "It was *like* [as if] she didn't understand" [lexical] → "She like didn't understand" [focus] → "She's like: 'I don't understand'" [focus quotative]).

   In conclusion, we have seen that nonstandard *like* plays a semantically rich and active role in some English-speakers' vocabularies. It follows systematic patterns of use and adheres to specific syntactic constraints. I question, however, along with Blyth et al. (1990) the need for a distinction between two functions of nonstandard *like*, for both may be shown to be related semantically as well as syntactically (through the use of *be*). The two roles of *like*, then, may simply be labeled FOCUS MARKER and FOCUS QUOTATIVE, that is, two different surface representations of the same underlying form.

REFERENCES

BLYTH, CARL, SIGRID RECKTENWALD, and JENNY WANG. 1990. "I'm like, 'Say what!' A New Quotative in American Narrative Discourse." *American Speech* 65.215-27.
BUTTERS, RONALD R. 1980. "Narrative *go* 'say'." *American Speech* 55.304-7.
FERRARA, KATHLEEN and BARBARA BELL. 1995. "Sociolinguistic Variation and Discourse Function of Constructed Dialogue Introducers: The Case of *Be + Like*." *American Speech*. 70.265-90.
MEEHAN, TERESA. 1991. "It's Like, 'What's Happening in the Evolution of *Like*?' A Theory of Grammaticalization." *Kansas Working Papers in Linguistics* 16.37-51.
ROMAINE, SUZANNE, and DEBORAH LANGE. 1991. "The Use of *like* as a Marker of Reported Speech and Thought: A Case of Grammaticalization in Progress." *American Speech* 16.227-79.
SCHIFFRIN, DEBORAH. 1987. *Discourse markers*. Cambridge: Cambridge University Press.
SCHOURUP, LAWRENCE. 1982. "Quoting with *go*, 'say'." *American Speech* 57.148-9.
TANNEN, DEBORAH. 1986. "Introducing Constructed Dialogue in Greek and American Conversational and Literary Narrative." *Direct and Indirect Speech*, Ed. Florian Coulmas, 311-32. New York: de Gruyter.
UNDERHILL, ROBERT. 1988. "*Like* Is, Like, Focus." *American Speech* 63.234-46.