

## INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.
2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.
3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again -- beginning below the first row and continuing on until complete.
4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.
5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

### University Microfilms International

300 North Zeeb Road  
Ann Arbor, Michigan 48106 USA  
St. John's Road, Tyler's Green  
High Wycombe, Bucks, England HP10 8HR

7902547

JANDA, ANITA MARY  
THE LINGUISTIC ANALYSIS OF THE HONEY BEE'S  
DANCE LANGUAGE.

CITY UNIVERSITY OF NEW YORK, PH.D., 1978

University  
Microfilms  
International

300 N. ZEEB ROAD, ANN ARBOR, MI 48106

© 1978

ANITA MARY JANDA

ALL RIGHTS RESERVED

THE LINGUISTIC ANALYSIS OF  
THE HONEY BEE'S DANCE LANGUAGE

by

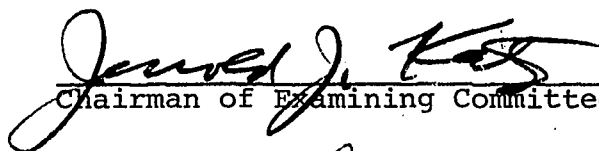
ANITA JANDA

A dissertation submitted to the Graduate  
Faculty in Linguistics in partial fulfill-  
ment of the requirements for the degree of  
Doctor of Philosophy, the City University  
of New York.

1978

This manuscript has been read and accepted for the Graduate Faculty in Linguistics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

September 25, 1978  
date

  
Chairman of Examining Committee

September 25, 1978  
date

  
Executive Officer

D. Terence Langendoen

Virginia V. Valian

Supervisory Committee

The City University of New York

## Abstract

### THE LINGUISTIC ANALYSIS OF THE HONEY BEE'S DANCE LANGUAGE

by

Anita Janda

Adviser: Professor Jerrold J. Katz

Now that the dance language has been experimentally established as a bona fide animal communication system, it is its linguistic status which is in question. The popular view of the dance language as a simple stimulus-response system, iconically coding the dancer's flight to the goal fails to so much as hint at the rich structural substrate of the behavior.

It is shown how published accounts of the various ways in which the communicative correlation is perturbed by circumstances mandate a grammatical description of considerable complexity--at best, a finite state device introducing three levels of labelling (i.e., mimicking a context-free phrase structure grammar:  $S \rightarrow \dots A \dots$ ,  $A \rightarrow \dots B \dots$ ,  $B \rightarrow \dots C \dots$ ) will be required. With the investigation of dance linguistic universals, however, the attempt to place the dance language in the Chomsky hierarchy, as requiring a grammatical description in finite state, phrase structure, or

transformational terms, entirely collapses.

This does not follow from any sort of species chauvinism "arguing" against the legitimacy of grammar construction, but rather is a mathematical consequence of the best definition of dance language grammaticality. A grammar is proposed which accounts for just which aspects of the behavior are susceptible to genetic variation, being independent of other factors, and which operate as a genetic package interracially persistent in the honey bee. In addition, this grammar explains why there are different dance types, why these are identically ordered by the various dance "dialects" via reference to goal distance, why  $30^\circ$  is the maximal angular "error" during a site advertisement, why there are direction-indicating dances which do not code distance, why the dances have the shapes they do.

The grammar consists of a set of semantic representations, each of which is a search area of fixed radius about a visited site, whose elements are paired with their syntactic codings by a geometric (not algebraic) averaging interpretive rule. The production and comprehension strategies suggested for implementing the grammar are identical, entailing that dancers and dance-recruited foragers contribute the same degree of error in actual communication. The grammar itself is proven incompatible with any of the linguistically motivated lexicons of countable (i.e., finite or denumerable) size.

The dance language thus has a nondenumerable lexicon, precluding placement on the Chomsky hierarchy of formal systems--no grammar specifying the proper ways of rewriting an initial start-symbol S is an appropriate analysis of the bee's system. The dance language is syntactically different in kind rather than degree from human language.

## Acknowledgments

Dr. Virginia Valian introduced me to the dance language and has (so far) managed not to tell me she's sorry about it. I have to congratulate myself on my choice of committee: Profs. Jerry Katz, Terry Langendoen and Virginia have been demanding but supportive throughout. The only one who's spent more time listening to me drone on about honey bees than they have is my sister Joan, but then, she's a relative.

I owe a special debt of gratitude to Drs. Mary Gallagher, Donald Griffin, R. M. R. Hall, Elliott Mendelson, Amy Myers, Alex Orenstein, John Ritter, and Joel Stemple; each of them gave me thoughtful encouragement at a time in my life when it mattered terribly to me. And again my committee should be mentioned, who (not so coincidentally) were responsible for making my years in graduate school stimulating and rewarding.

I am delighted to have this opportunity to thank Diana Kirigin who, out of pity for the feeble-fingered, actually typed the first draft for me; she knows I couldn't have done it without her, and now you know it, too.

Finally, I thank my parents, whose many sacrifices made me the first Doctor in the house.

## TABLE OF CONTENTS

1.	Introduction . . . . .	1
	Notes . . . . .	12
2.	Preliminary Observations: Focus on Performance .	13
	2.1 "Dance-Like" Behavior . . . . .	17
	2.2 Dancing . . . . .	19
	2.3 Dance Types . . . . .	22
	2.4 Distance Coding . . . . .	24
	2.4.1 Distance vs. Energy Drain . . . . .	34
	2.4.2 Distance vs. Flight Range Percentage.	41
	2.4.3 Determining the Linguistic Coding of Distance . . . . .	43
	2.4.4 Transport Experiments . . . . .	44
	2.4.5 The Averaging Hypothesis: Dance Production . . . . .	46
	2.5 Directional Coding . . . . .	49
	2.5.1 Detour Experiments . . . . .	52
	2.5.2 The Averaging Hypothesis Revisited: Directional Indication in a Crosswind	54
	2.6 Dancing Duration . . . . .	61
	2.7 A Production Strategy . . . . .	63
	2.8 A Perceptual Strategy . . . . .	65
	2.9 The Averaging Hypothesis: Dance Comprehen- sion . . . . .	65

2.10	Dancing Duration and Message Variation. . .	67
2.11	Light-Dependent and Residual "Misdirection" . . . . .	70
2.12	Disoriented Dancing . . . . .	77
2.13	The "Dance Language" Controversy . . . . .	81
2.14	Summary: Arguing from Performance to Competence . . . . .	87
	Notes . . . . .	94
3.	The Finite State Approach . . . . .	99
3.1	Dances Indicating Distance and Direction: The Carniolan Honey Bee . . . . .	99
3.1.1	The First Constituent of the Dance . . . . .	99
3.1.2	The Lexicon . . . . .	103
3.1.3	S-Dance . . . . .	106
3.1.4	Tailwagging Dance . . . . .	106
3.1.5	Transitional Dances: Y-Shaped . . . . .	108
3.2	The Grammar Thus Far . . . . .	111
3.3	Dances Indicating Distance and Direction: The Italian Honey Bee . . . . .	115
3.3.1	Transitional Dances: Sickle-Shaped. . . . .	116
3.4	The Round Dance . . . . .	118
	Notes . . . . .	124
4.	Alternative Grammatical Accounts . . . . .	127
4.1	Structural Criticisms of the Finite State Account . . . . .	133
4.1.1	Observational Adequacy . . . . .	133

4.1.2 Descriptive Adequacy . . . . .	163
4.1.3 "Explanatory" Adequacy . . . . .	191
4.2 Lexical Criticisms of the Finite State Account . . . . .	202
4.3 Syntactically Different in Kind Rather than Degree . . . . .	220
4.4 The Search Area Hypothesis . . . . .	242
Notes . . . . .	273
5. Implications, Experimental and Linguistic . . . . .	285
Notes . . . . .	301
Bibliography . . . . .	302

TABLE OF EXAMPLES

Example	Page
(1), (2)	22
(3)	55
(4)	56
(5)	57
(6)	66
(7)	71
(8)	76
(9)	103
(10)	104
(11), (12)	107
(13)	108
(14), (15)	109
(16)	110
(17)	112
(18)	113
(19), (20), (21), (22)	117
(23), (24)	118
(25), (26)	121
(27), (28)	123
(29)	127
(30), (31), (32), (33)	128
(34), (35)	129

(36)	130
(37)	136
(38)	137
(39)	145
(40)	146
(41), (42)	151
(43), (44)	152
(45)	153
(46), (47), (48), (49)	155
(50), (51)	157
(52), (53), (54), (55), (56)	158
(57), (58)	159
(59)	160
(60), (61)	161
(62)	162
(63)	170
(64)	172
(65), (66), (67)	175
(68)	176
(69), (70)	179
(71), (72), (73)	183
(74)	193
(75)	202
(76)	203
(77), (78)	205
(79)	206

(80)	217
(81)	218
(82)	249
(83), (84)	251
(85), (86)	253
(87)	255
(88), (89), (90)	257
(91), (92)	258
(93), (94)	259
(95)	262
(96)	263
(97), (98)	264
(99), (100)	268
(101), (102)	271

## INTRODUCTION

The issue of the possible evolutionary continuity of linguistic capacity is typically viewed from the perspective provided by an extensive list (explicit or implicit) of linguistic desiderata of an atheoretical nature, grounded in "common sense" observations. Among the more prominent attempts of this sort are Hockett 1960, Hockett and Altmann 1968, Gardner and Gardner 1971, Fouts and Couch 1976, Marler 1970, Nottebohm 1975, and Bronowski 1967. The criteria thus presumed operative act either to linguistically distinguish the human achievement or irrevocably link it to some other system. Virtually no attempt is ever made to independently analyze the semantic, syntactic properties of the nonverbal system prior to, or indeed after, the comparison with human vocal language. Once the decision is made that the system is remarkably inferior/similar to human vocal language, the nonverbal system itself is generally dismissed, its intrinsic linguistic interest ignored.

The emphasis here is upon determining the syntactic and semantic structures which best explain the patterning of the observed "utterances," as the most justified approach to the linguistic evaluation of the system. The

criteria properties of human language are a matter of debate--debate, however, as to what is the best linguistic theory. The criteria properties of human language are quite simply those entailed by the correct linguistic theory.

The conditioning design of the chimpanzee work incorporates a serious stumbling block with respect to justifying the linguistic interpretation of the behavior so painstakingly inculcated. Most recent attempts at linguistic overviews of the chimpanzee studies understandably, therefore, sidestep the knotty grammatical issue, targetting instead the blatant differences in the acquisition process (e.g., Mistler-Lachman and Lachman 1974) or the extent of the semantic capacity attained (e.g., Katz 1976).<sup>1</sup>

It was the desire to eliminate the confounding effects of the teach-the-chimpanzee-X experiments, where the demonstration of statistically significant levels of comprehension is all too often presumed to indicate learning of equal linguistic merit, which motivated the initial choice of the honey bee. The honey bee is both more difficult to anthropomorphize than the chimpanzee and far more extensively studied with respect to communicative capacity than any species other than man.

The dance language of the honey bee (Apis mellifera) is generally described (see, for example, von Frisch 1967a,

Lindauer 1967) as a genetically fixed coding of the distance and direction from the hive to the goal, the goal varying with the needs of the hive. The direction is given with respect to the sun's azimuth when the dance is done on a horizontal surface under the open sky (i.e., during swarming, when the goal is a new nesting site), and with respect to the upward direction when done, as usual, on a vertical comb inside the dark hive. Depending on how far away (distance gauged largely in terms of the sugar consumed on the outward flight from the hive) the goal is, the forager selects from her repertoire an appropriate dance form; each dance type is appropriate to a range of distance values. If the dance form features a wagging phase, during which the dancer wags her abdomen from side to side while simultaneously emitting a buzzing sound, it is the duration of this phase which codes the distance. The precise wagging-phase-to-distance correlation varies among honey bee races, as does their flight range, as a rough function of body size. Dances without a wagging phase can only give a distance range. All dance forms with straight runs, with or without superimposed wagging, code direction.

The analysis presented here grew out of an attempt (Janda 1977) to focus on the question of what power<sup>2</sup> grammar would be required to reveal the structural properties of the dance language of the honey bee. It will be claimed

that the difference between at least this species-specific information-coding system and human language is syntactically one of kind rather than degree (as is also concluded in Haldane and Spurway 1954:263; Hockett 1960:428; Chomsky 1975b:301). Such a conclusion needs to be substantiated, both by argumentation and, as it is here, by the specification of a structure-revealing account of the bee's system.

In recent years the major concern in the biological literature has been the issue of whether this well-defined behavioral repertoire which undeniably exhibits the proper correlations (is information-coding, has the property Hockett and Altmann 1968:63 labels "semanticity," maintains a symbol-meaning tie referential in nature) is in fact an animal communication system.

Many social and non-social animals "dance" excitedly after finding food, transform orientation angles from light to gravity, and convert distances travelled into durations of vibrations or sound production, and yet are not thought to communicate anything by these manoeuvres. Unsurprisingly, the question finally arose whether the evidence is really convincing that honey bees are uniquely endowed with a language. For example, although cricket songs correlate rather well with temperature, they are not thought to communicate that information to other crickets. . . . The critical question, then, is not whether distance and direction correlations exist in the honey bee dance, but whether this abstract information is communicated to recruits. (Gould 1975b:24)

The so-called dance language controversy, then, is concerned with whether the information coded in the dance, accessible to the human observer, is in fact accessed by (or even accessible to) the observing bees who follow the

dancer so enthusiastically. Dance following, it is argued (see, for example, Wenner 1967, 1974; Wenner, Wells, and Johnson 1969; Johnson 1967b; Johnson and Wenner 1966, 1970), may result solely in the "communication" of olfactory information. Recruitment to goals may be accomplished by olfactory and visual cues at the site.

While this issue is in fact addressed and the major experimental work in this area outlined in sec. 2.13 below, it is not at all clear that the outcome of this controversy is crucial to a linguistic investigation.

The standard Chomskyan conception of linguistics as a branch of theoretical psychology holds that those structural characteristics which recur uniformly across human languages are attributable to innate characteristics of the mind, features of man's "language acquisition device," and thereby features of the hypotheses available for the language learner to entertain in (unconsciously) constructing a grammar for a given corpus of linguistic data. Such universals are a reflection of human conceptual organization; in fact, linguistic universals, together with those uncovered by similarly rationalist cross-cultural scientific investigations into other areas of human knowledge, in effect constitute human mental organization. Hence the commitment to the claim that Language of a species-specific achievement is a natural, if not logically necessary, conclusion in the absence of accepted conceptual

non-linguistic universals. It then becomes a matter of theoretical interest to discover for any accepted conceptual universal whether it in fact cuts across species boundaries; it can be of only secondary interest to ascertain whether in cutting across species boundaries, it is found in similar cognitive domains.

The question under consideration here is whether transformational complexity, perhaps the least-questioned linguistic universal, may be adduced for the honey bee's dance language,<sup>3</sup> initially identified as a cognitive system through its information-coding aspect (e.g., von Frisch 1967a). It is maintained that, within the context of the standard Chomskyan view, this structural issue is of far greater significance than the demonstration that the bee's dance system is like Language not only in its information-coding role but also in its communicative function (Gould 1975a, 1975b, 1975c, 1976; Schricker 1974a).

The major concern here then is the demonstration that the dance language is amenable to a digital grammatical description only under an interpretation of the data which violates structurally significant features of the system. The claim at that point is that a grammar specifying the proper ways of rewriting an initial start-symbol *S* is not an appropriate analysis of the bee's system.

As will be shown, the popular view of the dance language as a simple stimulus-response system, iconically

coding the dancer's flight to the goal fails to so much as hint at the rich structural substrate of the behavior. The need for a sharply drawn, theoretically principled distinction between grammatical and extra-grammatical phenomena emerges quite early in the investigation and, not surprisingly, poses serious methodological problems. Among the problems apparent from even this very brief description of the dance behavior are (a) the question of whether at the grammatical level it is distance or sugar consumption (energy drain) which is linguistically encoded; (b) whether the linguistic coding of that parameter is the duration of the wagging or the duration of the buzzing; and (c) whether the discovery of racial dialects based on the differing wagging-phase-duration-to-distance correlations maintained by different races of honey bee is not an artifact of the distance/energy drain semantics assumed; reinterpreting the coding as giving, not an absolute figure, but rather the proportion of the flight range to be traversed, suggests that the dialects are nothing more than the result of the physiologically determined use of the same linguistic system by creatures of differing body size. The more fundamental question of the psychological weight of a grammatical hypothesis for this species is frankly outside the scope of this study.

The obvious ploy is to base the grammatical/extra-grammatical distinction on that between the properties of

a system and the properties of the exploitation of that system, a distinction basic to the Chomskyan conception of linguistics. While this gives a coherent, if somewhat vague research strategy for grammatical analysis, the usual mentalistic interpretation of the distinction (which is arguably correct for psychologically weighting the grammars of human language) is somewhat less plausible in this dance language case. So: "[w]e thus make a fundamental distinction between competence (the speaker-hearer's knowledge of his language) and performance (the actual use of language in concrete situations)." (Chomsky 1965:4). The grammar, viewed as a theoretical construct no more problematic than the physicist's idealized frictionless surface, is an account of competence, the idealized product of the language acquisition process. "Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly" (Chomsky 1965:3). Separate performance models of production and comprehension, not part of the grammar, with explicitly limited access to the grammar, account for "such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest (random or characteristic) in applying the ideal speaker-hearer's knowledge of the language" (Chomsky 1965:3).

With regard to human language, then, this distinction

is drawn mentalistically<sup>4</sup> on a knowledge/use dimension. It is simply not obviously appropriate to speak of the honey bee as having grammatical "knowledge." In what follows, the grammatical/extra-grammatical distinction is often (loosely) referred to in competence/performance terms. The grammatical residue which is left after siphoning off the (largely obviously perceptual) system exploitation phenomena is not easily labelled. The conceptual, cognitive, "knowledge" cast of the grammatical residue may not survive further investigation. By dance competence, then, is meant the subject matter which the dance grammar is a theory of; by dance performance, the functioning of that competence in the real world context. References to conceptual/perceptual parameters in the data are necessarily impressionistic at this very early stage of grammatical investigation. It must be understood that the analytical hypotheses entertained and evaluated are independent of the psychological interpretation of the grammar.

The analog analysis which is advanced, based on a superficially mysterious "error" pattern long noted in the literature, provides an explanation for the syntactic variety in dance types, which is perhaps the most puzzling aspect of the system from an evolutionary point of view. This explanation in turn hinges on the positing of a semantic interpretation rather far removed from the traditional distance/direction description outlined above.

The empirical implications of the analysis have not as yet been tested experimentally; indeed, none of the published data appears to bear on it, the dance language controversy having focussed research on performance questions singularly irrelevant to an analysis of competence, such as drug-induced disturbances in dance production or perception, translation problems posed by different "dialects," and attempts at provoking dance information/paralinguistic information disagreements, etc. The competence model proposed here (see sec. 4.4) is a first attempt at a characterization of the dance language's structural properties, a characterization designed to explicate the production data elicited in relatively natural circumstances. It is expected that some such account is a prerequisite for a thoroughgoing analysis of actual communicative interactions as a product of performance contributions by both the dancer and the observers.

Accordingly, the present paper is divided into five sections. This first section sets the problem and the approach taken in the analysis squarely within the context of theoretical linguistics. Sec. 2 presents some of the more salient features of the dance behavior, providing a relatively informal account of the notion of "a grammatical dance," incidentally putting the dance language controversy into linguistic perspective. This done, sec. 3 gives a finite state account of the language, which

enterprise is criticized structurally in sec. 4.1, lexically in sec. 4.2. Sec. 4.3 attempts to give substance to the claim that the dance language, given the criticisms of sec. 4.2, is syntactically different in kind rather than degree from human language--an attempt pursued in sec. 4.4, which treats the structural implications of the "error" pattern alluded to on p. 9, reanalyzing the data in analog terms. The concluding sec. 5 is largely concerned with the implications of the structural issues raised here for future experimental work as well as purely linguistic research.

## Notes

## SECTION 1. INTRODUCTION

1. There are some notable exceptions to this practice, among them Limber 1977.

2. To meet the goal of descriptive adequacy. "A grammar is descriptively adequate if it strongly generates the correct set of structural descriptions" (Chomsky 1965: 60). This is in contrast to the more limited goal of designing a grammar which is observationally adequate, weakly generating (enumerating) the sentences of the language without, however, any concern for the accuracy of the structure thereby assigned. For a discussion of these concepts and related issues see, for example, Chomsky 1965: 60-62, Bach 1974:179-204; and Levelt 1974b:16-41. The power of the grammar is determined by the mathematical properties of the rule types used. See sec. 4, introduction.

3. Language is being used here in the technical sense of a set of strings formed by concatenation on an alphabet. No claim as to the comparability of the structure of this set with human language is intended. That issue is treated below in sec. 3-4, especially secs. 4.3-4.4.

4. See Valian 1976 for a critical discussion of current positions with regard to the competence/performance distinction and its place in linguistic theory. The position presented here is that of the standard theory as described in, for example, Chomsky 1965 and Katz 1972.

### Preliminary Observations: Focus on Performance





The grammatical characterization of the honey bee's dance language is the topic of secs. 3-4, and indeed, the focus of this investigation. Unlike comparable analyses of human languages, however, this one features a lengthy critical discussion, this sec. 2, of the evidence (actual and potential) for a particular informal characterization of the notion "grammatical dance."

This necessary preliminary to the comparison of grammatical alternatives is normally quite brief--brief because there is an approximation to a generally accepted phonetic theory to serve as the vocabulary for representing the data of a given human language L and, even more importantly, there are obvious means of access to the grammatical intuitions of a native speaker of L, viz. the possibility of the linguist's acquiring sufficient mastery in L to be able to question the (possibly monolingual) L-speaker; the task is further simplified if English-L bilinguals can be located. A grammatical analysis of a human gestural language such as American Sign Language will be delayed by the need to explain and justify a visual analog of phonetic theory (see, for example, Stokoe,


Casterline, and Croneberg 1965, Bellugi and Klima 1975) while studies of child language are delayed by the need for a principled translation from performance data to competence idealizations (see Limber 1977). An investigation into a nonhuman system inevitably has both difficulties.


This section deals with a variety of problems arising at this preliminary stage to the syntactic analysis. What follows is an outline of the conclusions drawn in this section as to the characterization of the dance language.<sup>1</sup>

The genetically fixed behavioral repertoire which is the honey bee's dance language is identified as (an idealization of) those bursts of activity which typically occur in the midst of the massed foragers and code information structurally dependent upon the physical form of the exhibition. Uniform reports of unsuccessful attempts at decoding a dance form eliminates the form from the behavioral repertoire: "dance-like," such utterances are (apparently, at least) not dances. Two dialects are

generally reported; these share the round (  ), tailwagging transitional (  ), tailwagging (  ), and S-dances (  ),<sup>2</sup> differing with respect to

the shape of the transition from round to tailwagging transitional: the Italian bee exhibits a sickle formation

( 

shift ( 

Under the general rubric of "distance coding," arguments are presented (a) identifying the durations of wagging and sound production as potentially two different competence codings of the distance, (b) advancing the hypothesis that sugar consumption/energy drain is, as von Frisch 1967a suggests, a performance measure of the distance competence variable,<sup>3</sup> and (c) suggesting that sensory inputs from both the outbound and homebound flights routinely function to determine the distance value encoded (the averaging hypothesis).

The precise mechanism for directional coding varies with the physical context of the dance itself, as indicated in sec. 1. On a horizontal surface under the open sky, straight (plus or minus wagging) dance phases are run so that their average points directly toward the goal; orientation utilizes the solar azimuth. On a vertical surface in diffuse light that same angle (the angle between the azimuth of the sun and that of the

goal) is danced with respect to the upward direction: negative geotaxis substitutes for positive phototaxis. The averaging hypothesis is maintained for directional coding in production.

Generally speaking, the goal's location is specified more than once by the dancer, who is trailed by a variable number of potential recruits. Averaging is generally agreed to be a part of the comprehension (but not production) model. Strategies are outlined both for alerting foragers likely to be roused to a goal at a particular location relative to the hive (due to von Frisch 1967a) and for retaining a position as a dance follower for several circuits (due to Gould 1975b). Message "repetitions" are altered to accommodate solar motion even when the motion must be calculated by a dancer in the dark hive, to whom the sun is not visible. This aspect of the dance language, and this aspect only, is known to be learned.

Directional "noise" has two primary sources, apart from solar motion: competition between by-light and by-gravity modes of directional indication (light-dependent misdirection) and the "attractiveness" of the directly upward, downward, leftward, rightward orientations on the vertical comb (residual misdirection). These are judged to be purely performance phenomena. Truly disoriented dancing, in which directional indications vary wildly within a site advertisement, arises for a variety

of practical reasons, perhaps most obviously when the sun is at the perceptual zenith and the dance context is a horizontal surface bathed in sunlight.

A grammatical dance is then an instance of one of these dance forms, coding some degree of locative information about a site previously visited by the dancer. All such information is given relative to the dancing's site, using the solar compass. The round dance codes a distance range, nonwagging transitional dances additionally code a direction, and all wagging dances code both distance and direction. It is hypothesized in conclusion that long stretches of dancing are series of dances, simply adjacent as in a monologue and not syntactically conjoined, and that measures of the goal's quality such as the dancer's vivacity (von Frisch 1967a) which may, as Gould 1975b suggests, be quantitatively measured as the lateral extent of the wagging movements, be viewed as paralinguistic.

This position as to the informal characterization of a grammatical dance is justified in the pages below.

## 2.1 "Dance-Like" Behavior

As was argued in sec. 1, the behavioral repertoire of the honey bee is linguistically significant only insofar as it is information-coding. Consequently, this analysis of the dance language purposely excludes what may be described as "dance-like" behavior--the jostling run,

the buzzing run, the grooming (shaking) "dance," the jerking "dance," the trembling "dance," none of which has been shown to be information-coding, though they may signify readiness to engage in a particular form of behavior. Thus, the jostling run indicates the readiness to dance: "[f]or the most part dancing takes place only after several successful flights, but jostling occurs even on the first successful return home" (von Frisch 1967a:278), while the buzzing run serves "as the signal by means of which scout bees . . . induce the swarm cluster to decamp" (von Frisch 1967a:279). Whether or not this behavior is communicative, it appears to correlate with the performer's own readiness to depart. The addition, in the light of further evidence, of either or both of these to the information-coding dance language is ultimately immaterial, unless some crucial structural complexity is additionally discovered.

The grooming (shaking) "dance" is not clearly as impressive as these (see Allen 1959a, 1959b; Gahl 1975), despite its description as "a shaking movement of the bees whereby others sitting in the vicinity are called out to clean certain parts of their body that cannot be reached by their own grooming devices" (von Frisch 1967a:280), given that no distinction is drawn there between behavior analogous to going down for the third time in view of the lifeguard and that like crying for help. The grooming

"dance" is therefore, in the absence of any argument to the contrary, assumed to be a nonlinguistic indication of the readiness to be groomed, as drowning is a nonlinguistic indication of the readiness to be saved. The jerking "dance" has no known significance (von Frisch 1967a:282); the trembling "dance" is a product of such diverse calamities as old age, rough handling by the experimenter, or the addition of any of a chemically wide variety of poisons to the food (von Frisch 1967a:282-83). None of these "dances" plays a role in what follows.

## 2.2 Dancing

What is under consideration here and uncritically labelled a dance (as opposed to, say, a fidget<sup>4</sup>) essentially maintains a distance-direction correlation with a location able to supply one of the hive's most pressing needs. Thus the dance may indicate the presence of nectar, water (spread on the comb it evaporates to cool an overheating hive), pollen, resin, or even a new nesting site if swarming is imminent or perhaps already in progress. The needs of the hive are made known to its members in part by means of tropholaxis, the regular foodsharing process, which serves to identify the hive's need with that of each forager; the constant foodsharing keeps the members of the hive approximately equally hungry, thirsty, satisfied (von Frisch 1967a:266; Gould 1975b:16). An unresolved issue

centers on the question of whether the dance is "contagious," i.e., whether a dance attender might begin to advertise a site on, so to speak, the word of the dancer, without having first visited the site herself (Griffin pers. comm.).

It is an open question whether the quality of the goal, given by the dancer's "vivacity" (which may be quantified in terms of the lateral extent of her wagging movements in the dance--see Gould 1975b:16) is thereby linguistically encoded. It may be best viewed as paralinguistic, analogous perhaps to vocal volume, pitch, etc. as measures of the speaker's emotional state.

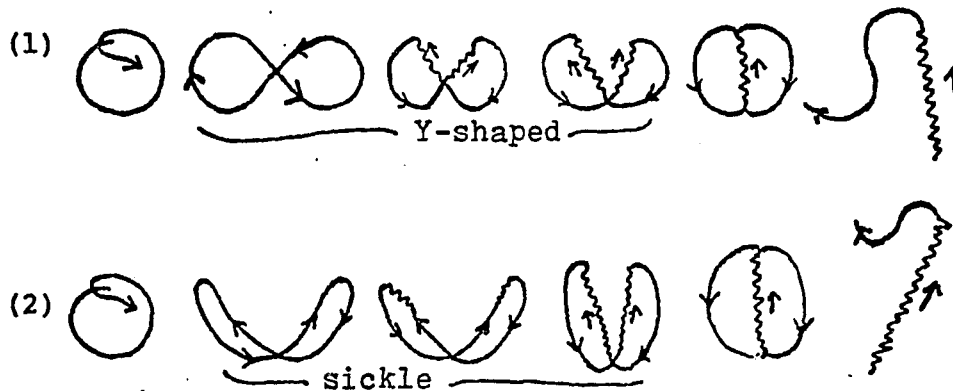
The linguistic status of the spasmodic "dance" is also unclear. Characteristically produced for goals very near the hive, the spasmodic "dance" consists of very short, apparently directed tail-wagging spasms, separated by irregular non-wagging runs interspersed with halts for food distribution (von Frisch 1967a:279). This "dance" is ignored in what follows due to the dearth of information with regard to the behavior; it is not discussed, for example, whether the directional correlation here is as accurate as that of the various clear cases of dance types, whether there is a distance correlation (and, if so, how its accuracy compares to that of the clear cases), whether the different honey bee races' and species' production use of the spasmodic "dance" patterns as does that

of the other dances with regard to the distance range indicated.


Interestingly, not all successful foraging flights provoke dancing. Indeed Gould 1976:212 describes the dance behavior as rare under normal circumstances. The more obvious factors to play a role in triggering dancing are the degree to which the site can serve to supply the hive's needs (so the sweeter the food the likelier to provoke dancing in a food-deprived hive, the less sweet the likelier to provide dancing in an overheating hive<sup>5</sup>), the size of the supply ("Even very sweet sugar water releases no dances if it is scanty"--von Frisch 1967a:238), the magnitude of the hive's need (more dancing in a starving hive), and the scent level of the site (slightly scented sites stimulate dancing while very strongly scented sites, easily identifiable olfactorily, inhibit dancing--in this regard, see sec. 2.13). In short, "whether dancing occurs at all, and how vivaciously and how persistently, depends on many factors that significantly regulate the relation between supply and demand" (von Frisch 1967a:236). Presumably, these and (possibly) other factors (von Frisch 1967a:236-56) may be assigned relative weights and the dance probability calculated for any particular constellation of circumstances.

### 2.3 Dance Types

Given a set of dance-provoking circumstances, the successful forager<sup>6</sup> enters the hive and proceeds to its dance floor (von Frisch 1967a:36-39, 138-39), where she traces one of the patterns in (1) or (2), the choice of






(1) and (2) redrawn after von Frisch 1967a:59,61

pattern being roughly determined by the distance of the hive from the site whose location is being coded: the farther away the goal, the more likely a pattern to the right will be selected. The  portions in (1) and (2) refer to stretches of the dance during which the bee energetically wags the lower portion of her body from side to side at an approximately constant rate; these tail-wagging dance phases normally coincide with the production of sound bursts via the flight musculature (Esch 1961, reported in von Frisch 1967a:58).

The set of patterns in (1) is characteristic of the Carniolan bee (*Apis mellifera carnica* Pollm.) while that in (2) is characteristic of the Italian bee

(A. mellifera ligustica Spin.) (von Frisch 1967a:294).

While both produce round dances (  ) for goals close to the hive, tailwagging (  ) dances at reasonably large distances, and S-dances (  ) for distances near the foraging limit, there are two features which serve to differentiate them, these being the shape of the transitional dances (Carniolan bees displaying a Y-shaped transition from round to tail-wagging, Italian bees a sickle formation) and the distance range assigned by each race to the dance types they share.




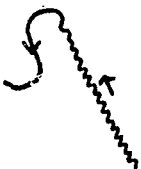
The Carniolan bee has a somewhat greater foraging range than the Italian bee (von Frisch 1967a:65-66) and uses each dance type for a correspondingly greater distance range. According to Boch 1957 (reported in von Frisch 1967a:295, Fig. 253), the Carniolan bee uses the round dance for distances up to approximately 16m. from the hive as compared to the Italian bee's approximately 8m.; tail-wagging begins for the Carniolan bee at about 85m., for the Italian bee at about 35m., with sickle dancing expected for distances between perhaps 9-22m., distance values not included are transitional stages, i.e. the choice of dance type is not even vaguely fixed for those values but fluctuates a good deal about the nearest dance types, perhaps combining them even within a single long site

advertisement (for a message is typically repeated several times at least). A final word of caution: these figures should be treated as approximations only and will vary across hives and indeed, to some extent, across individuals (Gould, pers. comm.), accounting for references to particularly "sluggish" or "lively" dancers.


The dependence of the distance code on the foraging range is quite reasonable in that, as von Frisch 1967a:125 notes in comparing it across species (the Carniolan bee with a flight range of perhaps 3 km. and the dwarf honey bee (A. florea F.) with a flight range of only 300-400m.), it "implies a precise indication of distance within the principal flight range, and no imprecision in the data until distances are reached that are scarcely of practical significance." Although the exact formula for converting to distance varies with the species, and indeed race, of honey bee, it remains as a general principle that the longer this wagging phase of the dance, the further away the goal is from the hive.


#### 2.4 Distance Coding

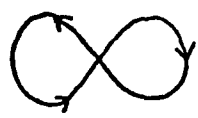
All dances featuring a wagging run, from the tail-

wagging transitional (  ), (  ),  
 through the tailwagging (  ) and S-dances (  ),


provide a measure of the distance from the hive (or swarm

cluster) to the goal. The round(  ) and non-

wagging transitional dances (again both the sickle 

and Y-shaped  racial variants) convey only a distance range within which the goal will be found.

Oddly enough, what von Frisch (1967a:104) describes as "the acoustically emphasized duration of wagging" is only one of many distance indications in the dance. Some degree of positive correlation with distance is provided by each of the following: the duration of the curvilinear return run (and therefore, the duration of a circuit, a circuit being a wagging run together with its curvilinear

return, so:  ), the length of the wagging run on the comb, and the number of wagging movements per wagging run (see von Frisch 1967a:97-104). It is there argued that it is wagging duration which is the definitive index of distance in the dance language. The arguments given are essentially eliminative in nature, criticizing the other candidates for this position. They are moreover not altogether convincing if, as seems likely, the claim is to be interpreted as drawing a linguistic/paralinguistic distinction with respect to the distance parameter of the

dance.

Gould 1975a:171 criticizes von Frisch's conclusion, pointing out that an approach based on dance production data necessarily fails to distinguish the wagging duration and sound production duration correlations, given that in "normal" (attention-getting?<sup>7</sup>) dances, wagging and sound production are coincident. In other words, von Frisch's "acoustically emphasized duration of wagging" might equally well have been described as the "wagging-emphasized duration of sound production." Gould's demonstration that the dance language is communicative as well as information-coding leads naturally enough to the position (adopted from von Frisch 1967a:104-106) that the resolution of this problem will depend on dance perception data (Gould 1975a:171), specifically that the solution pends the construction of an artificial bee acceptable to the hive.<sup>8</sup> The experimenter could then manipulate the different distance indications independently, creating a statistically significant degree of conflict between (say) the wagging and sound duration channels which is tolerated by the dance attenders; the recruit distribution pattern should then reveal which distance indication is used, i.e. which is the linguistic encoding of distance. In the event that the dance attenders will not in fact tolerate a statistically significant degree of conflict (ignoring such dances, for instance) it is to be inferred that the

linguistic encoding of distance is not unique.

The problem with this line of argument is that there is no reason to suppose that potential recruits will prefer linguistic to paralinguistic information in such a case. Would a man directing someone to walk three blocks north while pointing to the east succeed in sending his addressee north? If not, does this demonstrate that his pointing is the linguistic aspect of the interchange?

Indeed, were linguistic information generally preferred, the dance language controversy would never have arisen. The fact remains, however, that in a wide variety of circumstances, when both linguistic and nonlinguistic information are provided, recruitment takes place on a nonlinguistic basis. There are probably more examples of this phenomenon than of anything else in the honey bee dance language literature.

Von Frisch 1967a:22-23, for example, warns the would-be experimenter of two distinct olfactory dangers:

The odor of shellac clings to bees recently marked [i.e. numbered so the experimenter can focus on individuals] with alcoholic lacquers. Therefore the newcomers aroused by them go hunting for the odor of shellac--unless some stronger odor has been given to the feeding place--just as under normal circumstances they would seek the scent of the kind of flower that had been visited . . . and hence when they reach the feeding place they are more interested in the stand with the paint vials than in the feeding dish. Also the bees may in time become familiar with the individual odor of a regular experimenter at a feeding place. Then during a pause in feeding the bees may find the experimenter even if he is 100 or 200m.

away and well hidden, so that he is visited stubbornly by 'his' numbered bees.

Johnson, Wells, and Wenner, attempting to control for olfactory and visual cues in the field, found in a long series of experiments (discussed in sec. 2.13 below) that the preference for the feeding station to which the (presumably) dancing foragers were led vanished; paralinguistic information was preferred by recruits. In a similar vein, Lindauer 1967:126-30, reporting results of von Frisch and Lindauer 1954, notes that foragers trained to a given direction marked with a line of trees and then presented with a conflict between the (remembered) solar angle and a (new) line of trees (done by transporting the hive to a new location visually, olfactorily similar to its former site), search for the feeding station along the line of trees, rather than utilizing the solar angle which (again, presumably) would still be criterial for the danced directional indication.

Dance perception data can then be seen to play a distinctly minor role with respect to this issue, the model (dummy) bee no role at all, even if, for example, a noisy, non-wagging model proved acceptable. Dance perception data's only possible contribution here would lie in a demonstration that some of the distance correlations produced are literally not available to a dance attender due to the nature of her perceptual system. In which case, in order

for the analysis of the information-coding system not to preclude its interpretation as a communication system, this correlation should either be dropped from the competence model, and an explanation for it sought as a biological by-product of the linguistic correlation(s), or included in the competence model and analyzed as a crucial difference between the access separate production and perception performance systems have to that model.

The dance-production approach taken by von Frisch relies to some extent on an Ockham's razor simplicity criterion. Since the wide variety of completion curves in

the S-dance (e.g. , and all seem grammatical; see von Frisch 1967a:59, Fig. 51)

precludes their use as a measure of distance, the duration, length etc. of curvilinear return runs, and hence of the circuit, can not be the linguistic reflection of distance (von Frisch 1967a:102). The observed correlation in this case can only be a consequence of the true correlation(s) present in the single component common to all distance-indicating dances, the wagging run. Though the point is not made there, this construction of the case is supported by their relative efficacy as signals; wagging/sound production duration is the best measure of distance, 19% better than the duration of a circuit and as much as 352%

better than the duration of the curvilinear return run (von Frisch 1967a:102).

The simplicity criterion is then invoked on a psychological, as opposed to a purely linguistic, level to exclude the correlation with the number of wagging movements, which correlation is then apparently to be interpreted as a consequence of the true correlation (wagging/sound duration) together with the fairly constant rate of wagging within a swarm (von Frisch 1967a:57). The explanation offered for the dancer's monitoring of her dance should not credit her with the ability to count if other, psychologically simpler, accounts are consonant with the observable data.

Primarily, however, the concern is for the accuracy of each correlation: "What interests us is a comparison of the error (of the value as signals) of these elements" (von Frisch 1967a:100). It is assumed that the best correlation is the linguistic one; simplicity considerations arise secondarily to bolster the decision made from the numerical comparison. This is particularly obvious with regard to the elimination of the length of the wagging run as a distance measure, a correlation 170% worse than that of wagging/sound production duration (von Frisch 1967a:102). Simplicity is regarded as crucial only with respect to ruling out the number-of-waggles correlation,<sup>9</sup> which proves to be only 29% less good a signal than that of

wagging/sound production duration: "This difference is not so great as to exclude the number of waggles as a usable signal" (von Frisch 1967a:102).

Now while the demonstration that various features (the frequency of the wagging movements, the frequency of the sound produced, and the frequency and temporal structure of the vibratory episodes; see von Frisch 1967a: 98-99) of the dance fail to correlate with the distance is certainly significant in isolating the linguistic index of distance, it is by no means clear that the best correlation is the linguistic one.

In the present case then, three candidates survive the simplicity argument to persist as possible linguistic encodings of the distance: the duration of sound production, the duration of wagging, and the length of the wagging run. As already discussed, wagging and sound production duration are inextricable indications of the distance, considerably better than that provided by measuring the length of the wagging run. Were this last correlation physically independent of the duration correlations, the issue would be undecidable, short of a demonstration that some of these are not perceptible to the dance attenders. However, the wagging duration is clearly related to the length of the wagging run on the dance floor via the rate of linear progress made by the dancer in producing the wagging run. An unvaried dancing

rate would render the length of the wagging run precisely as good an indication of distance as the two duration correlations. Therefore, either (a) variability in the dancing rate has impaired the linguistic correlation present in the wagging/sound production mechanism or (b) such variability has acted on the linguistic encoding of distance present in the length of the wagging run to improve the correlation with the duration of wagging/sound production. Such a conspiracy toward accuracy, though logically possible, would require considerable explanation. The hypothesis that the duration of wagging/sound production is the linguistic encoding of distance is then the simplest available and is therefore adopted here.

This interpretation is reported in von Frisch 1967a:102 as confirmed by von Frisch and Jander (1957) who report a degree of recruit accuracy that would be inexplicable if the definitive cue were the length of the wagging run. Note that this result is at all relevant only if it can be demonstrated that none of the recruit accuracy was due to nonlinguistic olfactory/visual cues. In that case, it still remains to be shown that the "definitive cue" is the linguistic encoding of distance. As has been shown, this is not necessarily or even generally the case. This is not to say that honey bee communication displays an interaction between competing symbolic systems,

only one of which is linguistic (as in the north/east example for human language discussed earlier). Rather, what is considered here is the possibility that that aspect of the honey bee's apparently single symbolic system known as the dance language which is accessed by the dance attender is a biological by-product of the linguistic correlation.

The linguistic encoding of distance is determined as the most accurate member of each independently executed set of correlations, these sets first filtered by simplicity considerations. It is apparently not unique.

Gould's (1975a:171) sum-of-the-squares method is a statistically more sophisticated variant of the general approach of von Frisch and Jander which presupposes not only the now-familiar "the definitive cue is the linguistic one" but also that the encoding and decoding of this cue is mediated by the same mechanism. This last assumption implies that the statistical accuracy with which the production system accesses the competence model matches the statistical accuracy with which the comprehension (perception) model does so, i.e., that the scatter in production,  $p$ , is the same as the scatter in comprehension,  $c$ . Since  $p^2+c^2$  gives the square of the scatter,  $r$ , in recruit distribution, this gives a value of  $r/\sqrt{2}$  for the scatter of the definitive cue in production. Thus the straightforward measure of  $r$  and the production scatter

p along each correlating dimension (as well as that produced by the computed sum-of-the-squares interaction of any combination of these) should in time reveal what the definitive cue(s) are. Statistically more sophisticated, this method of analysis is linguistically even less so, gratuitously assuming as it does that encoding and decoding are computational inverses of each other, operating moreover with equal accuracy.

#### 2.4.1 Distance vs. Energy Drain

The linguistic/paralinguistic issue is further clouded by experimental evidence (see von Frisch 1967a: 109-21 for a fuller discussion) suggesting that in performance the correlations just enumerated are not in fact with the absolute (map) distance from the hive or swarm cluster to the goal or even with the distance actually travelled by a wind- or obstacle-detoured dancer previous to her dance. Rather, the most exact correlation is with the energy drain imposed by the trip.

Burdening the dancer with either lead weights heavy enough (55mg.) to impede take-off or comparable drag-producing tinfoil flaps prior to the foraging flight acts to deform subsequent "distance" indications; such dances signal a greater "distance" than would those of an unencumbered forager. The same effect may be achieved by shortening the dancer's wings before the dance-provoking

flight or forcing her to make the trip on foot (done by training her to feed in a gradually lengthened runway originating at the hive's entrance, its roof being too low to permit flight). The last factor is quite the most impressive; walking a distance of 3-4 m. provokes a ratio of round to tail-wagging dances like that triggered by a flight of 50-100m. (Bisetzky 1957, reported in von Frisch 1967a:116). Finally, situating the feeding station so that the flight from the hive involves a battle with a headwind or a flight up a steep slope, similarly serves to increase the indicated "distance."

Since a tailwind on the way to the goal reduces the indicated "distance" while a headwind increases it, the simple observation that a tailwind on the outbound flight generally ensures a comparable headwind on the return flight, demonstrates both that the correlation is not with total flight time (if it were there would in general be no wind effect), and that the "distance" indications must favor either the outbound or the homebound flight. The fact that burdening the dancer with weights etc. increases "distance" indications as does a headwind ("wind burden") on the outbound flight, argues for the distance indications' favoring this outward leg rather than the homeward stretch of the journey. Then, since walking 4m. takes considerably longer than does flying 100m. and yet provokes the same sort of dance coding (see previous

paragraph), the correlations cannot be with the time spent on the outbound trip either (Lindauer 1967:88). Experiments by Heran (1956, reported in von Frisch 1967a:115) suggest that the correlation is with the energy drain, an interpretation supported by measurements of Scholze, Pichler and Heran (1964) demonstrating that "the sugar consumption of a bee traversing 3m. on foot corresponds to that of a bee covering a course of 55m. in unimpeded free flight" (von Frisch 1967a:116).

The calculation of wind retardation and advancement of flight speed for experiments where the wind's direction and velocity were known revealed that the outward flight was not fully determinative of the dance-expressed energy drain. Von Frisch (1967a:128) describes the effect of putting a wind burden on the outward flight: "The tempo of dancing lags about 30-40 percent behind what would be expected if only the outward flight were taken into account." Or again, "it is as if in reporting the distance the outward flight weighs more heavily than the flight homeward" (von Frisch 1967a:119). Since the load on the return flight varies with the quality of the food supply and often is heavier than the 55 mg. lead weight that so affected the "distance" indications (see p. 34), this is presumed to introduce a certain amount of noise into the dance (see von Frisch 1967a:119-21).

In this regard it is additionally important to

determine the relationship between the energy drain/sugar consumption correlation when the goal is nectar (which itself provides the fuel for the homeward flight) and when it is water, pollen, or a new nesting site (which will not only not provide the fuel for the homeward flight but, in the water and pollen cases, will provide further burden on the return flight). Variation based on properties of the goal, should it occur, is of an inherently different kind from variation due to meteorological factors. The dance attendee can be presumed to receive no direct information about meteorological conditions prior to her own search, when she will be presented with similar, though not identical, environmental factors. The distribution of food or (presumably) water samples during the dance, the bodily evidence of pollen or propolis collecting, and (by inference from either the absence of such evidence or from the overcrowded, oversupplied condition of the hive) the evidence of nesting site inspection, all serve to identify the nature of the goal, whether or not this information is in fact accessed by the observing bees.

If it is the case that the "distance" correlation varies with the sugar (and/or burden) quality of the goal, it must be decided whether this qualifies as a case of contextually resolved ambiguity. The alternative is the position that one correlation is semantic and the other

a pragmatically deformed version of the primary (semantic) one (see Katz and Langendoen 1976:10). For example, the nectar-foraging correlation might be primary, with a [-sweet enough] context deforming the energy drain correlation by a constant factor, say  $1 \frac{1}{4}$  times the semantic correlation. An additional [-burden] feature may introduce an additional multiplicative factor of .7. The fact that a context-dependent (in the supply and demand fashion described on p. 21) threshold of sweetness is necessary to provoke food-advertising dancing would argue for a binary sweetness feature. The variation in the burden acquired at the goal may well, as von Frisch supposed (see p. 36), be a source of noise, but it is the presence/absence of burden on the homebound flight which may be anticipated during the dance, not the size of the burden, accounting for its binary quality.

If the variation is in fact constant in this way, this approach is preferable to the contextually disambiguating solution, where the pragmatic system appeals to the value of the sweetness and/or burden feature to select among lexically specified alternatives; within-lexeme regularities across racial dialects are a clear indication of under-analysis. Finally, there would remain the problem of deciding which feature values ([+sweet enough], [+burden] for the nectar-foraging context taken as the semantic standard in the previous paragraph) are semantic.

This is only one of an enormous number of linguistic issues surfacing with the discovery that the "distance" parameters of the dance correlate better with the energy drain than with the distance. The most important of these issues is the question of whether in this case competence mirrors performance, with neither assigning a pure distance measure to the dance; or whether, on the contrary, von Frisch's title for the section of his book treating this topic, "How Does the Dancer Estimate the Distance?" is linguistically apt. The general principle is that the performance system is physical/perceptual while the competence system is cognitive/conceptual. If there is to be a cognitive/conceptual aspect to this correlation, it must be this distance relationship. In other words, insofar as this aspect of the system is symbolic at all, the von Frisch approach is reasonable: holding such (in this interpretation, contextual) factors as weight carried, air resistance encountered etc. constant, reveals a clear distance correlation. The simplistic alternative denies that this is a symbolic aspect of the system, explaining it as a simple sugar-cost, energy-drain response only accidentally, as it were, providing a measure of distance.

This is not the place to attempt a definition of what constitutes a symbolic coding of information--it may in any event prove impossible to determine a species-

independent analysis of that concept.<sup>10</sup> It is with this considerable caveat that the indicated competence/performance line is drawn. At the competence level, the dances with wagging runs encode distance information where the distance coded in a particular dance token is a product of the dancer's perceptual (performance) system, viz. the energy drain the inspection or foraging flight produced.

It is important in this regard to determine whether other perceptual cues (for example, optical monitoring) interact with the energy drain measure. Von Frisch 1967a: 109 reports that Heran (1955) demonstrated that flight velocity is regulated by optical observation of the terrain:

[Heran] suspended a bee by the thorax and made her fly in one spot. In the darkroom she could see an illuminated, patterned background rolling away beneath her, as though she herself were going forward over a motionless region. She reacted to an acceleration of the background movement, that is, to an increase in her apparent flight velocity, with a decrease in her own exertion in flight (measured by the amplitude of the wingbeat).

If the "distance" indications in such a bee's dance were affected by her visual observations it would tend to support the competence/performance distinction as drawn above, with distance as the proper semantic parameter translating into a perceptual matrix at the performance level, a matrix whose entries may be put into a competitive relationship as sketched.

It is the simplicity of the supposed distance-gauging which is responsible for the problem.

#### 2.4.2 Distance vs. Flight Range Percentage

Yet another possibility is that what is being used as a distance gauge is not the energy drain, not the quantity of sugar consumed during the previous flight, but rather the dancer's estimate of the amount of the honey stomach to be filled with "flight fuel." The introduction of a positive (fuel tank) as opposed to a negative (energy drain) performance correlate provides a natural alternative to the orthodox distance interpretation of this parameter of the dance. The different distance values across racial dialects (see p. 23) may be directly explained by the interpretation of this aspect of the dance as encoding, not distance per se, but something equally removed from the perceptual input, the percentage of the flight range to be traversed in seeking the goal. As already discussed, the interracial variation in flight range, tentatively accounted for in terms of the observed variation in body size, mirrors that in the distance indicated across races by a given wagging run. This pattern is now directly accounted for.

The fact that hives which are not racially pure show a linguistic and physiological hybridization to the same degree might not be best read as proof of a genetically

linked linguistic heritage,<sup>11</sup> but rather a physiologically, metabolically determined use of essentially the same linguistic system. Only the Y-shaped vs. sickle transitional stages would be left unexplained--the sole indicator of dialects within the dance language.

Note that the dialect translation issue has no bearing on the question of whether it is map distance or the percentage of flight range which is being coded. Both accounts will fail should interracial communication in experimentally mixed hives succeed.

Experimental work in this area of dialect translation (see Steche 1954, Boch 1957, and the discussion of their work in von Frisch 1967a:298-300 and Gould 1975b:15) has to date failed to settle this crucial issue. On the one hand, even such minor indications of misunderstanding as Steche 1954 reports seem likely to prove significant: olfactory/visual cues may well account for apparent cross-dialect comprehension. On the other hand, apparent cross-dialect misunderstanding may be explained as due to a lack of dancing in the hive or the particular recruit's having failed to attend occurring dances, in effect locating the control station food while on a scouting mission completely independent of the dancer's activities.

What is needed,, of course, is an experimental test of interracial dance interpretation in a context guaranteeing that the information being used by the observer was

garnered from the dancer's dance.<sup>12</sup> It must be emphasized that the discovery that interracial communication succeeds would act to undermine current interpretations of this parameter of the dance. Any analysis "explaining" this supposed communicative success by having the dance attendee assess the dancer's racial makeup and interpret the dance accordingly (i.e., making the distance adjustment called for by her own racial makeup) is prohibitively complex psychologically, although there is no obvious alternative.

The distance vs. flight-range-percentage (energy drain vs. fuel tank) debate would be rendered academic by the discovery that interracial communication succeeds, unaffected by the discovery that it fails. A variant of the unorthodox percentage-of-the-flight-range approach to competence is presented in sec. 4.4 below, and this possibility is otherwise ignored here in favor of the distance interpretation standard since von Frisch 1947.

#### 2.4.3 Determining the Linguistic Encoding of Distance

Given the hypothesis that at the competence level it is distance which is encoded, however great a contextual (wind, gradient, etc) interference there may be on a particular occasion, the criteria previously developed (see pp. 31-33) credit as linguistic the most accurate distance indication from each set of interdependently executed

dance correlations surviving a simplicity scything. In view of the energy drain gauging of the distance, however, the standard of accuracy used for comparing the various correlations must be revised.

If the distance traveled is gauged by the energy drain imposed by the flight, then the parameter of the dance which happened to correlate best with the absolute (map) distance from the hive to the goal in a particular experimental series may well not be the one that correlated best with the energy drain and so, not the linguistic encoding of distance. Neither are the results improved by examining the dance scatter for the different factors over various instances of a goal's being situated at a constant distance from the hive; the constant distance means nothing unless the effects of the wind, gradient etc. are counterbalanced. Paradoxically, the linguistic indication(s) of distance are not best hypothesized as each set's best correlate with the distance but with the energy drain.

#### 2.4.4 Transport Experiments

At the performance level, there remains the problem of identifying the perceptual input which is criterial in gauging the distance. In the so-called transport experiments (von Frisch 1967a:116-18), foragers were accustomed by repetition to a radical displacement of the feeding

station as they fed. In time, the tendency to reverse the pre-feeding flight in order to return to the hive was overcome and a separate return flight learned for that feeding station.<sup>13</sup> The distance correlations obtained did not favor the pre-feeding flight as is normally the case (see p. 36). Neither did the dancer sum the two independently mastered flights which are apparently needed to forage from the site. The dancer's unexpected but uniform response in the dilemma is to give the mean of the two values (von Frisch 1967a:117). A fully adequate performance account should permit the prediction of responses that are uniform, as this one is. The explanation given in von Frisch 1967a:128 that the return flight is now treated as an "outward flight, but one now in the direction of the hive" does not seem to explain anything.

What seems called for here is a transport experiment which adds a burden to the second, "return," leg of the journey--a burden which both was not encountered in any guise on the first leg of the flight and which would not be a burden were the second leg reversed. (For example, the "return" flight might begin at the base of a steep hill.) The question that arises is whether the distance response is the mean of the two flights actually flown (as von Frisch's comments above suggest) or the mean of the first leg and the (calculated) reversal of the second flight.<sup>14</sup> This possibility, unlike that envisioned by

von Frisch, would be relatively easy to accommodate, as witness the following scenario.

#### 2.4.5 The Averaging Hypothesis: Dance Production

The dancer gauges the distance from the hive to the goal by separately estimating the energy drain (sugar consumption cost) of the pre- and post-goal flights. Where a contextual factor (wind, gradient, etc.) occurs which is as great a disadvantage on one of these flights as it is an advantage on the other, its contribution to the pre-goal flight is more heavily weighted. The functional explanation for this maneuver is that it prevents the factor in question from being ignored altogether; the processing explanation is that whatever effect, helpful or impeding, that factor has on the pre-goal flight creates a certain expectancy with regard to the nature and magnitude of its effect on the post-goal flight, thus reducing its post-goal contribution to the calculation. Other factors (e.g. the differing direction and distance aspects of the two legs in transport experiments) are simply recast in an outbound mode as separate, albeit surprisingly disparate, estimates of the same value, averaged to give the dance-coded estimate. Although the crucial experiment has yet to be done, it should be pointed out that the averaging and the reversal aspects of this interpretation are not only known to be within the honey bee's capacity,<sup>15</sup> but

additionally to function for directional determination in transport experiments:

[T]he displaced bees indicated quite precisely the bisector of the angle between the direction from hive to place flown to and from the hive to the place of return. . . [T]he solar angle perceived on the return flight is transposed by  $180^\circ$  and is taken into account in indicating the direction in the sense of the outward flight. (von Frisch 1967a:170, reporting results of Otto [1959])

This general approach will admittedly have to be modified somewhat as regards the directional indications in such cases, to accommodate the observed weighting of the two directional inputs.<sup>16</sup> Perhaps the most puzzling aspect of directional coding is the observation that "when there is a displacement of  $180^\circ$  the position of the sun during the outward flight generally is determinant; only when the bee flies away from the sun on the return flight (and on the outward flight) does she signal the place returned from; thus the position of the sun during the return flight, transposed by  $180^\circ$ , then decides the direction signaled (see Otto 1959)" (von Frisch 1967a:171).

Now, when the transport is to a locale such that the dancer does not require training in order to execute the return flight (i.e. can find the hive quite easily the very first time), only the outward flight plays a role (Lindauer 1963 reported in von Frisch 1967a:172-73). This is taken to demonstrate that under normal conditions the outward flight alone determines the dance-coded signals,

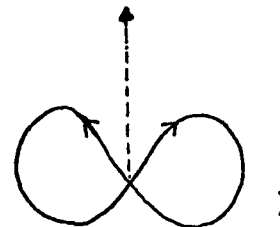
accounting for von Frisch's peculiar description of the trained return flight as an "outward flight, but one now in the direction of the hive." This phenomenon is however equally well explained by the current hypothesis augmented by the idea that the averaging process is generally selective, discarding outrageously outsized values. Apparently, training is required to confirm widely divergent estimates of the direction (and distance, presumably); without training, when the outward and homeward flights conflict the outbound leg is determinative.<sup>17</sup> This addition to the hypothesis is at least indirectly amenable to experimentation; it should not prove impossible to determine statistically whether the averaging over separate dance-codings of the site takes into account clearly outvoted values (see n. 15). The model bee project would prove helpful here.

It must be clearly understood that this issue arises at the performance level, concerning as it does the physiological, perceptual data utilized in determining the energy drain, then coded into the dance as distance. At the competence level, the linguistic encoding of distance is given by that element of each set of physiologically interdependent realizations to survive simplicity considerations which best correlates with the energy drain imposed by the pre-dance foraging flights, energy drain being the perceptual measure of distance. Insofar as the

effects of wind, gradient, etc. were counterbalanced, the finding (von Frisch 1967a:99, 101-102) that wagging duration and sound production duration are their set's best correlate with the distance is relevant to this determination. In that case, they are additionally the best correlates with the energy drain and thereby the best hypothesized linguistic encoding of distance.

## 2.5 Directional Coding


The linguistic encoding of direction poses no such problem. It remains an open question as to why there should be multiple distance correlations but only one directional signal (Gould 1975a: 171). The direction to the goal is given by the bisector of the wagging runs, superimposed in the tailwagging dance (von Frisch 1967a:141-42). The single wagging run of the S-dance also correlates with the direction to the goal.



Non-wagging Y-transitional dances ( ) correlate the direction with the bisector of the straight runs, but can give only a range as regards distance information. The bisector of what can be thought of as the axes

(see bold face arrows<sup>18</sup>) of transitional sickle dances similarly indicates the direction; again those without

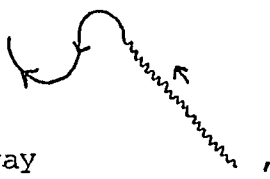



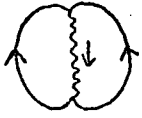
wagging can give only a distance range. The round  dance, lacking both straight and wagging runs, can accordingly signal neither distance nor direction but again only a distance range. It is a moot point whether the distance range indications are semantic or pragmatic, i.e. arise by a process-of-elimination "inference" from an examination of all the messages that were not sent despite their easy encoding.

When the dance takes place in daylight or on a horizontal surface (for the species and races in question, when the dance takes place in the swarm cluster), the resulting bisector points directly to the goal. The principle is apparently a simple one with respect to the tailwagging dance: the bee dances her wagging run in such a way that the sunlight<sup>19</sup> impinges most directly on ommatidia of the same dorsoventral circumference as it would on a direct flight from the hive to the goal (von Frisch 1967a:132, 136, 138). In the transitional dances of the Carniolan bee, it is presumably displaced by the same number of ommatidia on the second run as on the first, though in the opposite direction. Direction toward the sun is indicated by a wagging run toward the sun (positive phototaxis).

Directional gauging and encoding are, however, far more complex than this suggests, involving what may be thought of as a physiological translation mechanism when dancing takes place in the dark hive, as it ordinarily

does for this species. In this case, direction is indicated on the roughly vertical dance floor with respect to the pull of gravity: negative geotaxis is substituted for positive phototaxis so that direction toward the sun is indicated by a wagging run away from the earth (von Frisch 1967a:137). Thus a dance oriented on the dance floor the



 are on this page indicates a direction approximately 45° to the left of the sun, while


 indicates a direction approximately directly away from the sun.

As the translation analogy suggests, this is relatively easy to accommodate within the general framework sketched for directional encoding.

The coding of direction in terms of the azimuthal difference between the goal and the sun as viewed from the hive then may be expected to pose a problem in execution when the feeding station is at no horizontal distance from the hive.

Initial experiments suspended the hive against a cliff face and positioned the feeding station at some considerable height (i.e., at a distance expected to provoke tail-wagging) above or below it. It is reported (von Frisch 1967a:165-67) that dances directed into the

cliff led recruits up the cliff the proper distance, dances directed away from the cliff led the recruits the proper distance down the cliffs. Whether or not the dances were responsible for the recruits' success, well-oriented dances of the shape described demonstrate that the cliff experiments did not succeed in posing a performance problem. Orientation to the cliff face on an upward flight, away from the cliff face on a downward flight gave the vertical flights a "horizontal flavor."

Repeated with an open-work radio tower, it was found that suspending the food station 50m. above the hive provoked round dances in Italian foragers, who begin producing tail-wagging for (horizontal) distances of about 10m. (von Frisch 1967a:167-68). This result is offered there as evidence that there is no dance coding for elevation, but only horizontal distance and direction. As Gould 1975b:15 points out, if replicable it is inexplicable on the energy drain hypothesis, otherwise experimentally well-attested.

The general topic of performance problems in directional coding is further explored below (see pp. 77-81).

#### 2.5.1 Detour Experiments

The detour experiments of Bisetzky, von Frisch and Lindauer (see von Frisch 1967a: 173-85), however, reveal that when a direct flight from the hive to the goal proves impossible due to intervening obstacles, the directions flown on the separate legs of the pre-feeding flight are

integrated to give the air line direction from the hive to the goal indicated in the dance.<sup>20</sup> The averaging here is unlike that earlier described for the transport experiments where both the distance estimates and the directional estimates may be produced by averaging; here, directional estimates are averaged, while distance estimates are summed (von Frisch 1967a:176-77). The crucial difference lies in the nature of the estimates made; in the transport case, the conflicting estimates occur on the outbound and return flights and thus presumably reflect a single value for each parameter (and are therefore averaged), while in the detour case, the conflicting estimates are purely directional, occurring during the outbound flight, and indeed, averaging takes place along that parameter.

There is thus considerable evidence against the straightforward perceptual interpretation of directional encoding in terms of the sunlight intake of each ommatidium. In this regard, it is additionally noted (see von Frisch 1967a:186-90) that when the flight is undertaken in an appreciable crosswind, the bee orients herself into the wind, so that (just as in the detour experiments) the light may never impinge most directly on the ommatidia corresponding to the danced directional indication, which again gives the map-correct direction.

### 2.5.2 The Averaging Hypothesis Revisited--Directional Indication in a Crosswind

The fact that the forager heads into a crosswind in this fashion forms the basis of an argument given in von Frisch 1967a against the averaging hypothesis presented above (pp. 46-47) with regard to the transport experiments' data. The averaging hypothesis, it will be recalled, claims that the distance and direction values signalled in the dance were gauged by averaging the estimates of these values made on the outbound flight and (the calculated directional reversal of) those made on the homeward flight; the average generally favors the outbound estimates, heavily weights landmarked laps and, finally, ignores a widely divergent homebound estimate unless it is repeatedly confirmed (as by training). Von Frisch 1967a consistently repudiates this view maintaining, as indicated above, that although averaging certainly occurs along the different legs of the outbound flight (accounting for the air line indication in the detour experiments), the fact that some transport experiments (those where training to the return flight was not required) show no effect of the homeward flight, precludes any theory of outward-homebound integration under normal circumstances, i.e. where the homebound flight is the reversal of the outbound flight.

The simplicity argument levelled against the averaging hypothesis in the crosswind case (von Frisch 1967a:195) has two clauses. The first objects to an averaging analysis

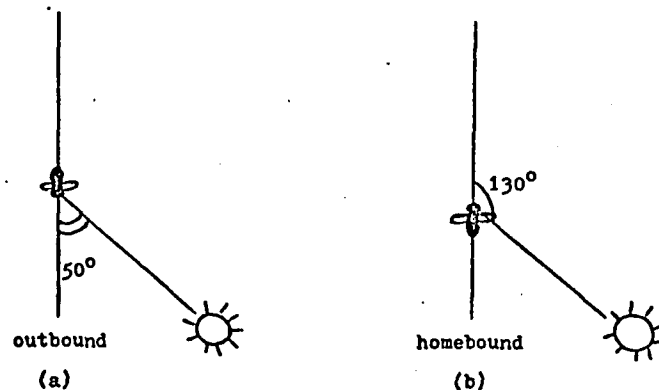
here, given the "capabilities that [the dancer] has displayed so unambiguously in the detour experiments." This is a simplicity razor that cuts both ways--given our reanalysis of the transport experiments using the averaging hypothesis, it is clearly simpler to suppose that the averaging capability recurs in the case of a detour or crosswind.

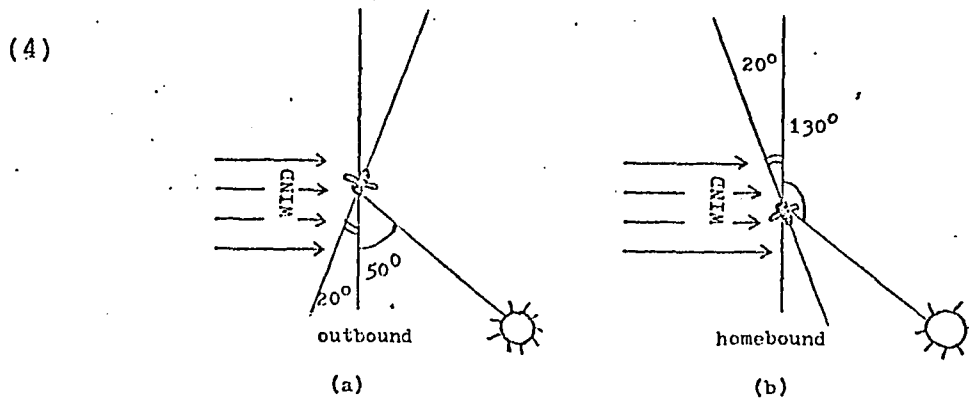
The second clause deals with the biological complexity of the averaging analysis in this crosswind case:

[w]hereas in calm air the transposition imagined can be interpreted relatively simply, . . . to the effect that on the return flight the bee shifts to the hinder region of the opposite eye the solar angle that she had before her on the outward flight, with the body in an oblique position . . . this simple transposition does not lead to the goal; rather the oblique position has to be included here in the calculation, by some sort of integration process.

The example given is of a crosswind from the left provoking a  $20^\circ$  orientation into the wind, transforming the situation sketched in (3) to that in (4):

(3)



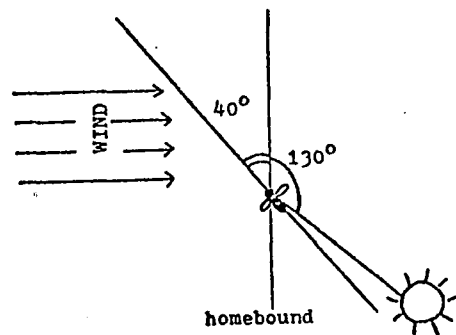


Now the averaging hypothesis has the  $50^\circ$  dance-coded value in (3) produced by averaging the outbound (3a) value of  $50^\circ$ , with  $180^\circ - 130^\circ = 50^\circ$ , the supplement of the home-ward (3b) value of  $130^\circ$ . According to von Frisch (1967a: 195), the averaging hypothesis is committed to producing the  $50^\circ$  dance-coded value in the crosswind case (4) by averaging the outbound (4a) value of  $50^\circ + 20^\circ = 70^\circ$ , with  $180^\circ - 150^\circ = 30^\circ$ , the supplement of the homeward (4b) value of  $130^\circ + 20^\circ = 150^\circ$ . The reason for this is clearly the non-calculating, non-computing view of averaging (or, more usually, integration) adopted from Jander 1957:194 (quoted in von Frisch 1967a:185): "By integration is understood the central nervous reduction of complex spatial and temporal patterns of sensory perceptions to one resulting datum." If averaging is central nervous system processing, then the actual sensory inputs are naturally viewed as determinative.

The second clause of von Frisch's criticism of the averaging hypothesis, that the integration is here too

complex to be true, is then reinforced by his observation that the averaging hypothesis as interpreted in the previous paragraph will in the crosswind case only average to the map-correct direction if the wind is reasonably constant in direction and intensity. To give an example, had the wind shifted while the bee in (4) was feeding so as to produce a  $40^\circ$  heading into the wind on the return flight, the proposed integration would have produced an incorrect dance-coded value of  $40^\circ$  by averaging the outbound  $70^\circ$  as in (4a) with the homebound  $10^\circ = 180^\circ - 170^\circ$  as in (5).

(5)



This of course assumes that the wind shift compensation is reasonably immediate, deduced for example from the sensory perception of the airstream (both antennae and eyebristles apparently play a role in this--see von Frisch 1967a:187). Greater discrepancies than that illustrated using (5) will be produced by more severe or more numerous wind shifts on dance-determining collecting flights.

Von Frisch's alternative to the averaging hypothesis

is as follows. The outward flight is fully determinative except where the return flight has been laboriously learned and thereby functions as an "outward flight, but one now in the direction of the hive" (von Frisch 1967a:128). Averaging is understood to occur only with respect to the legs of the outbound flight and therefore plays a crucial role only in detour experiments and those transport experiments in which the return flight is learned. The 50° dance-coded value in (4) is to be produced solely from the sensory data encountered during the outward (4a) flight, where the perceived solar angle is 70°. The 20° sensory correction for the wind effect is hypothesized to be due to the visual estimate of the lateral drift of the landscape rolling away beneath her (von Frisch 1967a:186). In support of this, it is noted that an appreciable crosswind will blow a bee off-course only in the absence of landmarks (Heran and Lindauer 1963, reported in von Frisch 1967a:186).

This account of the crosswind data systematically devalues von Frisch's attack on the averaging hypothesis. The major criticism of the averaging hypothesis had to do with its complexity, the fact that "the oblique position has to be included here in the calculation, by some sort of integration process." This can now be seen to be a necessary feature, not of the averaging hypothesis, but rather of von Frisch's own account. Only the "simpler"

von Frisch account forces a sensory correction of the perceived solar angle data; the averaging hypothesis can correct for wind effects using the perceived solar angle of the return flight, i.e. in (4) averaging  $70^\circ$  with  $30^\circ$  to produce the dance-coded  $50^\circ$ . In other words, from the point of view of the averaging hypothesis the oblique stance into the wind which is undeniably part of the bee's orientation mechanism need not play any sort of integrative role in determining the dance-coded direction; the dance-coded direction may be produced by simply averaging the perceived solar angle on the outbound and homebound flights. It is von Frisch's non-averaging hypothesis which makes the dance-coding a matter of correcting for the oblique stance imposed on her by the wind.

As von Frisch notes, this interpretation of the averaging hypothesis predicts that the dance-coding in wind shift cases will be incorrect, unlike his own account. Until such time as it is demonstrated that in such cases the dance-coded direction correctly corresponds to the air line, however, this cannot be considered a problem. Note that in the event that wind shifts do provoke errors, some version of the averaging hypothesis is called for--the von Frisch analysis can not adjust to accommodate such a possibility. The averaging hypothesis can, however, be revised to account for the opposite eventuality, viz., accurate air line dance-coding of direction in the face of rapid, severe

wind shifts during collecting. For, in that case, it must be supposed that the wind correction mechanism functions to filter the solar angle perception so that in, for example, (4), the 50° dance-coded value is produced by averaging the 50° outbound value sensorily calculated from the 70° solar angle perceived and the 20° lateral drift observed below (as it is calculated on von Frisch's account) with the 50° homebound value sensorily calculated from the 30° solar angle, due to directional reversal of the 150° solar angle perceived, and the 20° lateral drift. The averaging hypothesis is not, as von Frisch suggests, committed to determining the 50° value in (4) by averaging 70 and 30°; it is conceivably that these values are first wind-corrected to 50° and 50°, and only then averaged.

Even in this case, it seems clearly arguable that averaging (which is admittedly the only possible account of the dance-coded directional value in detour cases as well as both distance and directional values in the trained transport cases) is most simply hypothesized to operate at all times. The averaging is apparently constrained to require confirmation of widely divergent perceptual estimates (as by repetition, i.e. training) and quite possibly operates on lateral drift-corrected percepts. Neither of these caveats need be viewed as undermining the hypothesis whose alternative (i.e. von Frisch's account) requires lateral drift filtering in any event and restricts averaging to

detour and trained transport cases, in which the return is an "outward flight, but one now in the direction of the hive."

## 2.6 Dancing Duration

The message is typically repeated several times with minor and apparently insignificant variations in the signal, of which von Frisch (1967a:72) writes "[t]he bee is no precision machine." This creates a constituency problem in some cases, notably as regards the round dance where the reversals are strikingly irregular: "Between two reversals there are often one or two complete circles, but frequently only three-quarters or half of a circle. The dance may come to an end after one or two reversals but 20 and more reversals may succeed one another . . ." (von Frisch 1967a:29). A constituency issue is also raised by the following observation with regard to the tailwagging dance: "The semicircles over which the dancer returns approximately to the starting point after the completion of her tail-wagging run are as a rule traversed alternately to right and left, but exceptionally two right turns or two left turns may succeed one another" (von Frisch 1967a:59). By analogy with the tail-wagging transitional dance, however, whose semantic interpretation requires both a left and a right circuit to derive the directional bisector, it may be argued that this dance has both a left and a right circuit,

i.e., that problematic series featuring two successive right or left circuits (e.g., RLL, LLRR are problematic, LRRL is not) contain an acceptable but ungrammatical string. No such helpful analogy occurs for the round dance, which indicates neither distance nor direction to the goal.

Some of the variation described may legitimately be attributed to "inattention," while the cluttered condition of the dance floor ("There is never any dancing on an empty or sparsely occupied comb, but only in the midst of a throng of bees" (von Frisch 1967a:29)) is no doubt responsible for much that remains:

Just as with the round dance, the tail-wagging dancer is trailed after by other bees that, with extended antennae, seek to keep in touch with her and acquire information. . . . the event takes place in the midst of the massed throng, while the dancing group in its energetic rounds pushes aside the uninterested portion of the other hive members and thus makes itself sufficient open space. (von Frisch 1967a:59)

When the dancer has been foraging rather than inspecting prospective nesting sites, a dance follower may interrupt the dance with a squeak, receiving in return a honey stomach food sample from the dancing forager (see von Frisch 1967a: 29, 59-60; also see n.6). Still, references like the following to the general "slovenliness" of the dancing remain a problem: "Mostly the dancer does not remain precisely on a single spot. As she shifts across the comb the course traversed is somewhat deformed. Even in a single location the pattern may be repeated exactly or with less

regularity, one might say in a slovenly way" (von Frisch 1967a:61).

## 2.7 A Production Strategy

Von Frisch 1967a discusses what may be described as a site advertisement production strategy, designed to promote contact with those bees most likely to participate in the foraging from that site. The need for some such strategy for the dance language arises due to the following consideration: "Under normal conditions there are in a hive several simultaneously active groups of foragers visiting different kinds of flowers. If at times the bees of one such group find no food at the accustomed flowers they do not immediately transfer to another source" (von Frisch 1967a:30). It is then imperative that some mechanism exist to direct the dance, alerting those foragers predisposed to seek out the foraging place, maximizing recruitment. The production strategy consists in correlating the area of the dance floor used for a particular site advertisement with the location of that site relative to the hive.<sup>21</sup>

The clarity of the distance relation ("with increasing distance to the feeding places the dancing areas were moved stepwise upward and backward from the hive entrance (von Frisch 1967a:138)) raises the possibility that this is not a performance strategy at all but rather yet another aspect of the linguistic encoding of distance. While not

absolutely ruled out, this interpretation is unlikely given the nature and extent of the directional correlation between dancing area on the dance floor and site advertised: "the direction of their forthcoming dances . . . participates in determining the direction of their entry [onto the dance floor], insofar as that is possible and so long as they are not thereby deflected too far from the area of the comb where the food recipients are sitting, into other regions where they would not be freed of their burden" (von Frisch 1967a:141). In practice, this means that the directional correlation held only prior to dances indicating a roughly toward-the-sun (i.e. upward) direction, where indeed the scatter of entry direction was additionally reduced. In other words, the directional correlation is clearly a subordinate part of the alerting function of the dance behavior. A similar interpretation of the distance correlation as a performance strategy is therefore indicated. This two-parameter performance strategy is then mirrored by a converse comprehension strategy on the part of the potential recruits, who huddle on that part of the dance floor likely to serve as stomping ground for their group (group members are olfactorily identified-- see von Frisch 1967a:32).

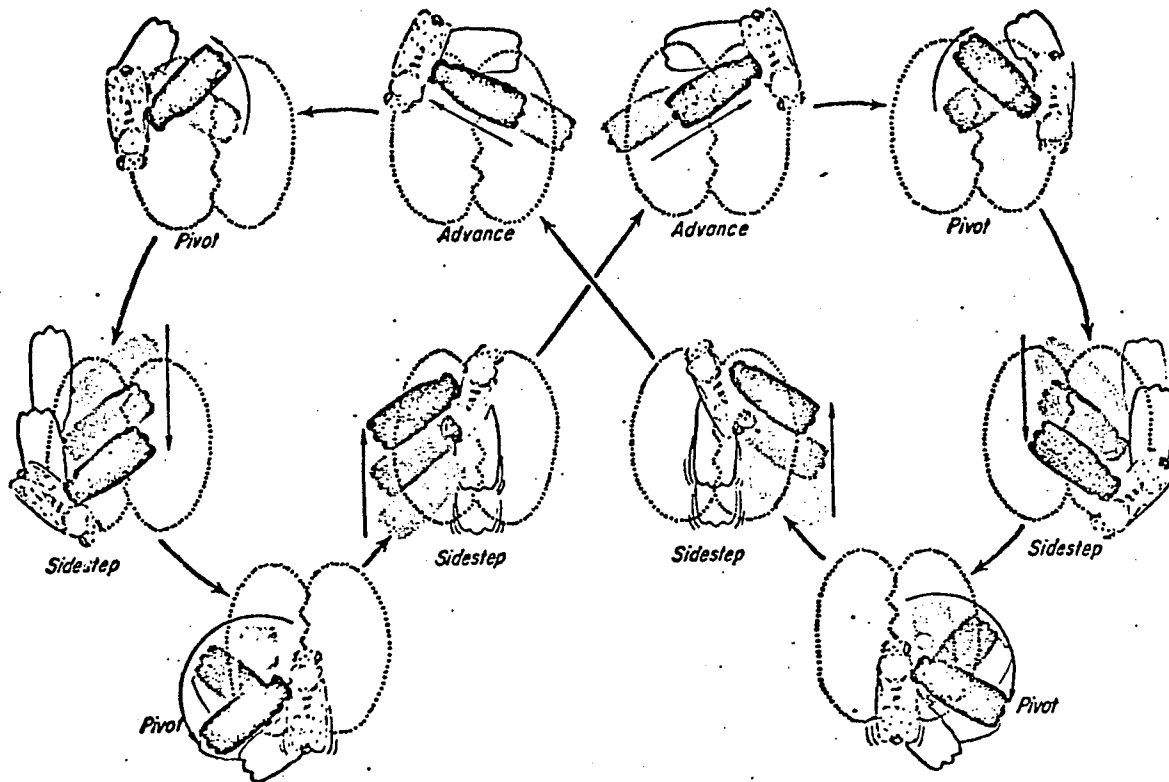
## 2.8 A Perceptual Strategy

A more interesting perceptual strategy is that sketched in Gould 1975b:110-11 for dance following which involves maintaining attendance at a relatively constant angle of about  $40^\circ$  to the dancer during the straight/wagging run, and a position primarily on the inside of the danced pattern the rest of the time: "[W]hen the dancer turns away at the beginning of the [curvilinear] return, the attender advances, thereby cutting off attenders which have been 'watching' from the other side" (Gould 1975b:111). This is the strategy of those dance followers who manage to maintain contact over an appreciable series of pattern traversals. There is substantial agreement on the number of cycles attended by successful recruits, with Esch and Bastian 1970, Johnson and Wenner 1970, Mautz 1971 all setting it at about six (see Gould 1975b:47-54, 172). The approximately  $40^\circ$  discrepancy between the bodies of the dancer and the dance follower is alternately to the left and to the right during the direction-giving phase of the dance using the strategy in (6), with the strategy failing when two left or right circuits occur in a row.

## 2.9 The Averaging Hypothesis: Dance Comprehension

Perceptual averaging or integration arises not only as a possible explanation of dance production, but also dance perception. Here, however, the crucial experiments

(6)



(Gould 1975b: 111, Figure 38)

have been run; averaging in dance perception is demonstrable from the accuracy with which recruits alerted by the dance (see sec. 2.13) seek the goal. Regardless of which cue is the definitive cue of distance in communication, it is not produced with sufficient accuracy to account for the statistical distribution of recruits without the hypothesized averaging (Gould 1975a:688-91; 1975b:170-73; 1975c: 171).<sup>22</sup>

## 2.10 Dancing Duration and Message Variation

There is considerable variation in the dance types produced during a stretch of dancing:

As a rule when the distance of the feeding place is 25m, a few short, definitely directed wagging runs are to be noted. At 50m, such runs may be seen with about half the dancers, and at 100m the tail-wagging dance is dominant; even there scattered round dances may still occur from time to time. Some degree of divergence of the wagging runs may be preserved at distances up to several hundred meters. (von Frisch 1967a:62, my italics)

Or, again, "[c]ounts made on photographs showed S-curves only after 3-4 percent of the wagging runs for food places as much as 1000m distant . . . after 14 percent for 2000m . . . and 33.3 percent after 4500m" (von Frisch 1967a:59). There is in addition some variation within the separate energy drain (distance) indications themselves (see von Frisch 1967a:59, 100-102; Gould 1975c:171). This variation cannot simply be a product of such factors as wind and gradient, since it recurs within what is generally treated as the repetition of a message and, what is more, to approximately the same extent (von Frisch 1967a:70-73).

This raises the question of whether the variation in distance indications is itself responsible for the production of different dance types within a long site advertisement. That is, in a mixed (borderline?) message repetition are the different distance indications conveyed by the dance type reserved in the grammar for such distances? To put it yet another way, do separate dance types share

wagging/sound duration values?

The directional variation within "marathon dances," which may last for several hours and, indeed, into the night (Lindauer 1967:95), presents a very different problem. In the course of such a dance, the sun's azimuth changes more or less drastically; the dancer, in the dark hive, increments her directional signal to compensate for the calculated solar motion typical of her hive's locale at that time of day (Lindauer 1967:115; von Frisch 1967a: 350-52). The honey bee's incrementing capacity is apparently further attested during the training of a forager to a food site within the experimental context: "one can . . . slide the table farther away, at first only a few centimeters at a time . . . Very gradually one can proceed by longer steps . . . the table should finally be advanced about 100m at a time. It is as though the bees had grasped the movement of the feeding place, for at times they are awaiting one at the next station" (von Frisch 1967a:117). Experimentation controlling for olfactory/visual cues is necessary to confirm this.

The aspect of the problem which is of linguistic interest is the fact that in a marathon dance compensating for solar motion, dances are seriated (conjoined?) which are apparently not all synonymous, combining any number of different directional signals. This issue is treated in some detail in sec. 4.1. Here it suffices to point

out that dances occurring in the swarming situation last an average of fifteen times longer than those of a dancing forager (von Frisch 1967a:269). This may be due in part to the reduced number of searching clues provided in this non-foraging and therefore non-food sample context.

Another factor which may play a role in this is the necessity for the swarm cluster to agree on a new nesting site within the time limit imposed by the size of the food supply appropriated on leaving the hive; generally speaking, it is only when all dances advertise the same site (von Frisch 1967a:269-73) that the dancers break into the buzzing runs which appear to signal departure. No such agreement occurs in the foraging situation.

The linguistic compensation for solar motion is the only aspect of the dance behavior which has been shown to be learned. Lindauer (1957, 1959, reported in von Frisch 1967a: 207-10, 364-65 and Lindauer 1967:121-26) raised a colony of "cellar bees" in an incubator, without access to a view of the sky. After being trained to a southern feeding place located (at that time of day) to the left of the sun for several afternoons, they were displaced (to prevent the use of visual landmarks) and observed to seek it to the left of the sun in the morning, in an easterly direction. It required some five afternoons of open-air flights to learn the solar compass. Even here there is evidence of an innate endowment--the capacity to calculate

the entire diurnal course of the sun from experience of just one small segment of it. As von Frisch 1967a:365 notes, this is confirmed by the rough accuracy<sup>23</sup> of overnight marathon dance angular progression.

There is some question as to whether there might not be a critical period for such learning. It may be significant that where naturally reared colonies were displaced across the equator, adaptation to the radically different course of solar motion required a length of time suggesting the birth of a new generation (Lindauer 1967: 118-19 reports a value of 43 days, noting that the bee's life span is some 4-6 weeks). On the other hand, it may simply be the case that such knowledge, once acquired, is not susceptible to modification, although its initial acquisition is not subject to any time or developmental pressures.

#### 2.11 Light-Dependent and Residual Misdirection

The directional indications on the vertical comb are subject to two perturbing influences labelled respectively light-dependent and residual "misdirection." Light-dependent misdirection arises when the dancer catches a glimpse of the blue sky, whose polarization pattern permits her to (perceptually) deduce the position of the sun; the competition between the two modes of orientation (positive phototaxis and negative geotaxis) results in her

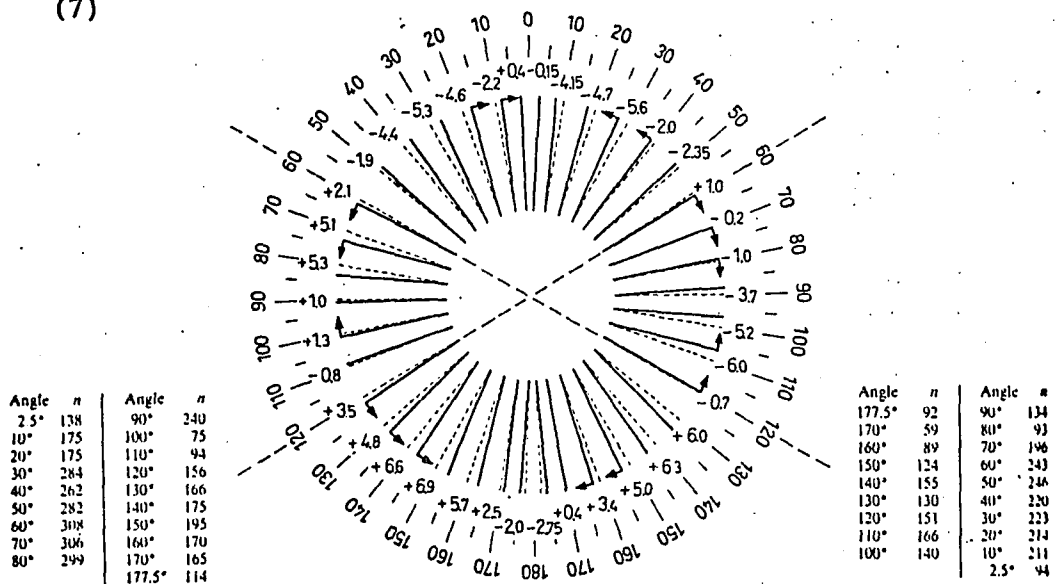
dancing the bisector of the two values (von Frisch 1967a: 197-204, 233-34).

All that matters is in what direction and at what angle the sun stands relative to the bee dancing on the vertical comb surface. That depends not only on the position of the sun, but also on the position of the comb surface in space and on the position of the bee in her tail-wagging dance (which differs, for instance, under otherwise identical conditions, on the two sides of the comb). (von Frisch 1967a:197)

In the event that the sun itself may be glimpsed, the orientation is to the light, the gravitational attraction playing no role (von Frisch 1967a:202-203).

Residual misdirection varies with dance direction (see (7)) and was initially thought to be due to organic

(7)



Over-all average values from all usable experiments regarding "misdirection" (33 experiments with 6759 measured dances), arranged according to the angle of dancing. The direction of dancing is to be followed counterclockwise, as are the numbers (n) of the individual values in the lists at left and right... dashed lines, expected directions of dancing; solid lines, observed directions of dancing, combined in means for each 10°. For each stage the amount of misdirection is given; its tendency to shift in the course of the experiment is shown by the arrows.

(von Frisch 1967a: 214, 216 Figs:210, 212)

translation problems encountered in converting direction with respect to light into direction with respect to gravity (von Frisch 1967a:217). However, it has been found to be a purely gravitational phenomenon; the tendency to err toward the vertical and (to a lesser extent) the horizontal on the vertical surface (see (7)) is not peculiar to linguistic behavior but recurs when the bee is trained to seek food at a given angle on a vertical surface (Markl 1966, reported in von Frisch 1967a:221-22). The vertical aspect of this deflection at any rate is a widespread phenomenon among arthropods, although the horizontal tendency in inaccuracy seems peculiar to the bee--see von Frisch 1967a:212-22. The deviation in the outward run to the food source in Markl's work agrees at all angles measured with the danced directional deviations (residual misdirection), which of course corresponds to the direction between the hive and goal.<sup>24</sup> Markl's explanation is that a species' inherited fundamental orientation (positive, negative, and transverse geotaxis for the bee) "is the original manner of orientation from which menotaxis, the setting of the body, on the basis of experience, at the chosen angle relative to the source of stimulation, is to be derived" (von Frisch 1967a:219-20), "that the geomenotactic orientation learned is influenced lastingly by the inherited fundamental orientation" (von Frisch 1967a:234).

Both misdirection effects are performance phenomena.


Indeed, the light-dependent misdirection averaging on the vertical is a special case of a more widespread weighted averaging on an oblique comb surface (not unusual during swarming). Provided the comb is at least  $5^\circ$  off the horizontal,  $5^\circ$  being the approximate physical limit on her gravity perception mechanism, the size of the by-gravity contribution is tied to the angle of inclination of the comb: the smaller the angle, i.e., the more vertical the comb, the greater the by-gravity contribution. At the same time, the size of the by-light contribution is affected by the cloud conditions, in the obvious way. (See von Frisch 1967a:146-49 for a fuller discussion.)

It is claimed that neither light-dependent nor residual misdirection misleads recruits, who are subject to the same influences as the dancer and so can compensate for their perturbing effects (von Frisch 1967a:203-204, 212-13). Such claims need to be confirmed with olfactory and visual clues eliminated,<sup>25</sup> given that demonstrating such a deflection in production does not by any means demonstrate an inverse, compensatory ability in comprehension as this suggests. What's more, as Gould (1975b:110) points out, the  $40^\circ$  discrepancy between dancer and dance follower (see (6)) itself guarantees that they are not identically affected. Thus perceptual decoding by a dance follower will not yield the dancer's pre-misdirected directional value but only an approximation to it.

The issue of whether misdirected dances misdirect recruits is of importance in determining the nature of the comprehension model, and indirectly of the production model as well. Consider for a moment what the given interpretation of light-dependent misdirection in production implies for the comprehension model, assuming that light-dependent misdirected dances are correctly interpreted.<sup>26</sup>

Suppose, that is, as von Frisch 1967a:201 does, that the dancer averages "the directions that would have corresponded to orientation by gravity and by light, respectively." In other words, that her danced directional value  $x$  on a comb from which the sky is visible corresponds to an averaging of  $s+\theta$  and  $\theta$ , where  $\theta$  is the value that would occur were the blue sky not visible (i.e. for the by-gravity orientation aspect), and  $s$  is the difference between the direction away from the earth and the dancer-calculated solar azimuth, for the by-light orientation aspect.<sup>27</sup> Or, alternatively, that  $x$  corresponds to an averaging of  $\phi$  and  $e+\phi$ , where  $\phi$  is the value that would occur were the sun visible (i.e. for the by-light orientation aspect) and  $e$  is the difference between the dancer-calculated solar azimuth and that away from the earth, for the by-gravity orientation aspect.<sup>28</sup> The basic idea behind von Frisch's account of light-dependent misdirection is that comprehension consists in the recovery from the danced value  $x$  (the average of the by-light and by-gravity

specifications of the goal) of an approximation to the pre-misdirected value--the value that would have been danced either in the absence of polarized light ( $\theta$ ) or the sight of the sun itself ( $\emptyset$ ). This forces a de-averaging capacity on the comprehension model, recovering  $\theta$  from a sense datum  $x$  averaging  $s+\theta$  with  $\theta$  and one specifying  $s$  (or, what amounts to the same thing, recovering  $\emptyset$  from a sense datum  $x$  averaging  $\emptyset$  with  $e+\emptyset$  and one specifying  $e$ ).

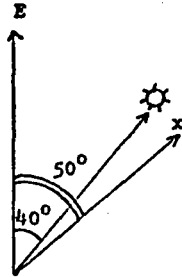
The alternative to von Frisch's account, almost painfully obvious, is to suppose that it is the two modes of orientation brought into competition which are averaged. The observer, like the dancer, averages these values. Her comprehension of the "misdirected" dance is accounted for without position additional ("de-averaging") conceptual apparatus; the danced angle  $x$  is read with respect to the appropriate axis ( $s/2$  or, equivalently,  $e/2$ ) just as it is when the sun is visible (  ) or when the light is diffuse (E). The two approaches are compared in (8).

Note that the von Frisch account provides two distinct but empirically indistinguishable models, while the alternative account, which is both psychologically simpler (lacking de-averaging) and makes a uniform account of all directional comprehensions possible, provides only one model.

The best hypothesis, then, for the comprehension (and so, production) systems is that polarized light

(8)

(a)



von Frisch

$$s = 40^\circ$$

$$x = 50^\circ$$

$$x = \frac{s + 2\theta}{2}$$

de-averaging

$$50 = \frac{40 + 2\theta}{2}$$

$$\theta = 30^\circ$$

direction with respect to E

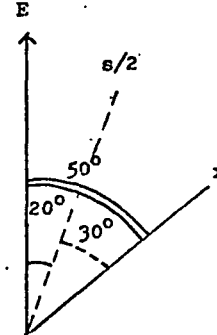
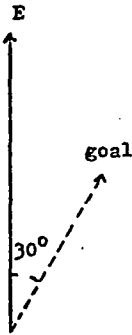
alternative

$$s = 40^\circ$$

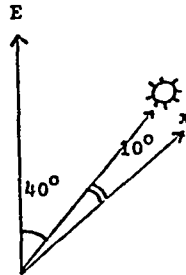
$$x = 50^\circ$$

$$s/2 = 20^\circ$$

vs. direction with respect to s/2



(b)



von Frisch

$$e = -40^\circ$$

$$x = 10^\circ$$

$$x = \frac{e + 2\beta}{2}$$

de-averaging

$$10 = \frac{-40 + 2\beta}{2}$$

$$\beta = 30^\circ$$

direction with respect to

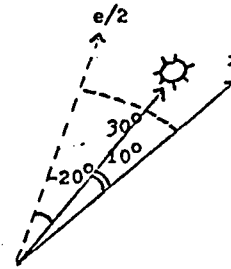
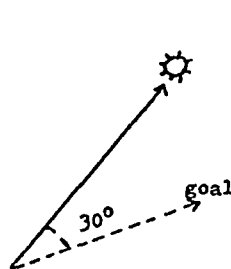
alternative





$$e = -40^\circ$$

$$x = 10^\circ$$

$$e/2 = -20^\circ$$

vs. direction with respect to e/2



triggers an averaging of the alternative axes, rather than an averaging of the alternative wagging runs themselves. When visible,  takes precedence over E as an axis; when the location of  must be perceptually deduced from the polarization pattern, its axis advantage is lost and  and E are on an approximately equal footing (weighted as described on p. 73, above), perceptually averaged by both dancer and observer (whose perceptual estimates of  disagree to some extent, as already discussed<sup>29</sup>).

Should recruits fail to correctly interpret (see n.26) misdirected dances, this issue will not arise. In that case, the comprehension model will not fully mirror the production model and it will therefore not be possible to argue from likely properties of the one to likely properties of the other.

## 2.12 Disoriented Dancing

The mode of directional coding on the dance floor requires for its proper functioning both that (a) the azimuthal difference  $\alpha$  between the sun and the goal be calculable from the perceptual data of the foraging flight, and that (b)  $\alpha$  be replicable with respect to the axis dictated by the dance's physical setting. Directionally disturbed dancing is then to be anticipated when (a<sub>1</sub>) the sun's position is not perceptually calculable, being hidden;

(a<sub>2</sub>) the sun's position is so near the zenith that its azimuth is not perceptually calculable; (a<sub>3</sub>) the goal's azimuth is not perceptually calculable; (b) the lighting in the vicinity is diffuse or lacking, calling for by-gravity directional coding, but the dance floor is so nearly horizontal that the dancer's gravity organs can not detect the upward direction with respect to which  $\infty$  is to be given.

Of the four anticipated sources of directional disturbances, perhaps (a<sub>2</sub>) has received the most attention. Although the lack of olfactory/visual cue controls makes much of the early work (e.g. New 1961) uninterpretable, it is reported that the incidence of dancing decreases as the sun nears the zenith (Lindauer 1967:91) and more importantly, that such dances as occur are "disoriented," exhibiting no directional agreement during a site advertisement (Lindauer 1967:92, von Frisch 1967a:163) when executed on the horizontal dance floor. New and New 1962 report that a small number of bees manage well-oriented dances in these circumstances when dancing the direction with respect to gravity, anticipating the future position of the sun as it would leave the zenith.

As regards (a<sub>1</sub>), it must be noted that the bees are "weatherwise," becoming discouraged by threatening weather in direct proportion to the distance between their goal and the hive (von Frisch 1967a:245). Lindauer 1967:93

further reports that bees are particularly sensitive to ultraviolet light (citing von Frisch, Lindauer and Schmeidler 1960), allowing for solar location even when it is quite overcast. As the cloud cover becomes impenetrable, direction is given with respect to gravity (von Frisch 1967a:202).

An (a<sub>3</sub>) example is provided by the radio tower experiments (von Frisch 1967a:167-68). These are purported to demonstrate that the elevation of the goal (and therefore presumably of the sun) plays no role in directional coding. What is significant, again, is the report that the response here was in terms of the round dance, although disoriented tail-wagging occurs in the other circumstances. The description of the disoriented dance notes that "not infrequently probing circular runs are interspersed among the dance figures" (von Frisch 1967a:147), so that this may be only a statistical difference, not significant linguistically at all.

Disoriented dancing is the uniform response to (b) (see von Frisch 1967a:146-47).

Von Frisch suggests that the response to situations of type (a) is the round dance, that to type (b) disoriented tailwagging: "Disoriented dances always take place if a dancing bee is unable to orient herself while performing the waggingrun . . . [I]f she cannot orient herself according to the sun on the flight to the goal . . . she

performs round dances in the hive" (von Frisch 1967a:134-35). The problem with this generalization, of course, is von Frisch's own report of disoriented wagging as the dance response in the dark hive to a near-zenith sun (von Frisch 1967a:163, result due to New and New 1962). A disoriented wagging response to a near-zenith sun is consonant with von Frisch's generalization only when the dance is executed in by-light circumstances--yet this response is also observed in the dark hive, i.e. by-gravity context. The report then of the round dance response in the radio tower experiments is not only problematic as regards the interpretation of tailwagging as the measure of distance or energy drain (see pp. 51-52) but also as regards the general topic of interference with directional coding.

The further question of whether disoriented dancing codes distance is also unsettled. Von Frisch's introductory discussion of case (b) concludes with the claim, credited to Esch (1964), that noise and circuit duration is distorted dances on the horizontal comb fail to vary with the distance, as they do normally (von Frisch 1967a:135). Yet von Frisch 1967a:153-55 reports a recruit pattern in such cases which suggests that distance information is coded. Now, while distance coding might prove to be dependent on the cause of the disorientation (as, for example, von Frisch's suggested (a) vs. (b) bifurcation) or even the physical context of dance execution

(by-light vs. by-gravity), these two results are not compatible. Either there was some olfactory or visual cue present at the recruit stations the proper distance from the hive or the salient distance cue in communication is something other than what Esch chose to measure. The question of distance coding must then be investigated with respect to dances disoriented as per (a<sub>2</sub>) and indeed (a<sub>3</sub>), should the round dance not be the response in that situation, as seems likely.

### 2.13 The "Dance Language" Controversy<sup>30</sup>

The most obvious challenge to the dance language as an animal communication system was predicated on the idea that recruitment might be accomplished by olfactory cues left behind at the site, viz., hive odor and Nasanov gland secretion. Various attempts (see, for example, Lindauer 1971) to control for the Nasanov secretion involved sealing the scent gland with shellac; unfortunately, shellac is also an olfactory attractant (to the point where shellac bottles distract foragers from food sources-- see von Frisch 1967a:22-23), rendering this a control for Nasanov gland secretion but not for scent labelling of the site by the forager. The fact that sites which provoke dancing garner a larger number of recruits may be ascribed to the dance's having a general arousing effect; it need convey no locative information to the dance follower.

Simplicity considerations then militate against the imputation of an information-coding communication system to the honey bee.

There followed a spate of articles which found that by providing encouragement of various kinds at the control stations, much of the preferential force in recruitment of the experimental (presumably dance-advertised) station vanishes. Johnson 1967b, for example, trained mutant foragers, strikingly different visually from those of the experimental hive, to visit the three control stations; the experimental station attracted no more of the experimental foragers than did the control stations. Wenner et al. 1969 demonstrated that sponge "bees" feeding at a site also provide landing encouragement.

No matter how many contexts are discovered and detailed (as by Johnson 1967a, 1967b; Johnson and Wenner 1966, 1970; Wenner 1967, 1974; Wenner et al. 1967, 1969) in which recruitment is accomplished by nonlinguistic cueing, this cannot establish that there is no dance cueing (as, for example, is claimed in Wenner et al. 1969). Olfactory conditioning can--and no doubt does, in the light of this work--play an enormous role in the ecology of foraging behavior, but the so-called dance language hypothesis can not be disproved by such work, but only called into question. The simplest conclusion, in the face of data showing that recruitment is fully accounted for

in nonlinguistic terms, is, to be sure, one treating the dance language as behavioral correlations serving no known function, but that conclusion is only falsifiable, not verifiable. The literature which it is then important to examine in detail is that which describes attempts at falsification of the anti-communication position.

In 1974, Gould, Henerey, and MacLeod disproved the Nasonov version of the olfactory hypothesis with an experimental design utilizing two food sources in different directions. After training, sugar solutions of different concentrations were offered at the two sites--nearly all of the ensuing dances advertised the same (sweeter) site, which attracted nearly all the recruits. The visual and olfactory controls designed to make the two stations equivalent were heroic: for example, more foragers were trained to the less attractive (control) station because that station attracted fewer forager visits and it is forager visits which give the site landing encouragement, scent marking, etc. Assuming the controls were sufficient, the fact that the experimental, dance-advertised station attracted nearly all the recruits strongly suggests that there is directional information transfer of some sort during dance following.

This experiment is unusual in that it controlled for the occurrence of dancing by the trained forager. In point of fact, the Wenner method of training foragers to

the experimental station (Wenner 1961), which utilizes a sugar solution heavily scented with the experimental scent, unlike that favored by von Frisch or Gould (Gary and Witherell 1971, a modification of that outlined in von Frisch 1967a), allows for a marked build-up of the experimental scent in the hive, in which context dancing is known to be rare and olfactory recruitment the norm (Wells and Wenner 1971).

Now, Gould et al. 1974, in demonstrating information transfer during dancing, still does not demonstrate that it is the dance information which is accessed. Any attempt to control for (visual and olfactory) nonlinguistic cueing must make assumptions as to what sensory aspects are important for bees. In the present case, locale odor is a possible nonlinguistic explanation for the recruitment pattern, as it is for all the experiments thus far discussed (including those outlined in von Frisch 1967b). The landscape which was chosen for its apparent uniformity might not be olfactorily uniform for the bees. This is noted by Gould et al., who describe what is needed to evaluate the dance language's communicative efficacy: "A conclusive experiment might be one in which the information contained in the dance of the forager is altered in such a way that recruits using this directional information would proceed to a location to which the dancing forager had never been" (Gould et al. 1974:553). If recruits are "misdirected"

then the dance language is in fact an animal communication system, however relevant the work of Wenner, Wells, and Johnson is with respect to honey bee recruitment in general.

There have been a number of attempts to cause a dancer to "lie" about where she's been. Goncalves 1969, for example, trained his forager to walk through a tube to the food; since distance is calculated by an energy drain measure, her report of the distance to the food should overestimate it from the point of view of the (flying) recruits. As Gould 1976:223-24 observes, the fact that recruits arrived preferentially in the direction of the experimental station whether or not the tube was long enough for the forager to have performed direction-indicating dances (Bisetzky 1957) strongly suggests that this result was due to nonlinguistic cueing of some kind. Considerable technical problems emerge with this design, as with the model bee. Schricker (1974a) and Gould (1974; 1975a, b, c) have independently produced workable "misdirection" techniques, demonstrating that the dance language is indeed communicative.

Schricker 1974a is an investigation into the effect of parathion poisoning on the distance-coding capacity of the dancer. Earlier studies had shown that sublethal doses of this substance have two significant effects on directional coding<sup>31</sup>: (a) that "[t]he basic form of the

dances performed by poisoned and non-poisoned bees was similar, except that the angles indicated by the poisoned bees did not change continuously with time, keeping pace with the sun's movement across the sky, but did so in a step-wise fashion" (Schricker and Stephen 1970:150); (b) that "parathion-treated bees dancing on a vertical comb were unable to transpose their visual signals in any direction except the cardinal compass points and their intersects" (Stephen and Schricker 1970:164). The fan experiments of Stephen and Schricker 1970 revealed a statistically significant shift in recruit distribution after poisoning the dancer with parathion: the recruits appear to have been "misdirected." A negative correlation between wind direction and recruit distribution further supports the view of the dance language as a communication system (see discussion, Stephen and Schricker 1970:163-64).

Schricker 1974a pursues the question of the hyperactivity effect of parathion noted in Schricker and Stephen 1970:148 ("the parathion-treated bees [fly] approximately 10% faster than the untreated") in terms of its effect on distance coding. The poisoned dancer's increased dance rhythm reduces the wagging duration time substantially, thus coding a smaller distance than would a non-poisoned dancer. Recruits were misled, demonstrating that the dance's distance information is, in some contexts at least, used by recruits. The dance language communicates

distance information.

Gould 1975b's "misdirection" experiment took advantage of earlier results (Schricker 1965, Lindauer and Schricker 1963) showing that the ocelli, the simple eyes lying between the two compound eyes, function as light receptors; when a bee's ocelli are covered with paint, she has a shorter working day, requires higher light levels for phototaxis. A dancer with painted ocelli will indicate direction geomenotactically in a hive with an artificial light source bright enough to reorient the untreated dance followers, for whom the direction is apparently being indicated with respect to the light. The dance directs recruits to a site the dancer has not visited, a site she can not give nonlinguistic cues for. Changing the relationship of the light to the vertical changes the site advertised. The statistical preference for the direction (Gould 1975a) and distance (Gould 1975c) advertised in the dance demonstrates that the dance language is indeed communicative, coding both distance and directional information which potential recruits do in fact access.

#### 2.14. Summary: Arguing from Performance to Competence

The dance language of the honey bee is a behavioral repertoire coding information as to the distance and direction from the hive to the goal (nectar, water, nesting site, etc. depending on the current needs of the hive). The

extent of the locative information coded is variable with some (i.e. non-wagging) dance types failing to give precise distance or even directional information (i.e. the round dance). Both the vivacity of the dance and the sugar burden qualities of the goal may prove to play a linguistic role as well. The details of the dance language vary across races and species of bee, specifically as regards the shapes of the transition from round to tailwagging dance; it is an open question whether the interracial variation in distance assigned a particular wagging run is linguistic as well, given that it correlates with the interracial variation in flight range, dependent in turn on body size. A bout of dancing typically gives the location of the goal at least several times.

The distance is calculated in terms of the energy (sugar consumption) cost of an attempted direct flight to the goal, but may be found to involve other perceptual data as well. Direction is calculated from the sunlight intake of the various ommatidia, possibly corrected using visual estimates of her lateral drift when there is a crosswind to be fought. Only the goal's azimuth, not its elevation, plays a role in the direction coded. Directional coding fails, giving rise to disoriented dances, when the sun nears the zenith.

Evidence is presented for the averaging hypothesis, according to which perceptual inputs of various sorts are

integrated by the nervous system both previous to dancing and during dance comprehension, where it is generally agreed that approximately six inputs are averaged. The averaging hypothesis incorporates a filtering device which discards perceptual data widely divergent from previous experience (so that training is required to permit the integration of outward and homeward flights of very different lengths) and prefers the outbound flight, perhaps because many of the features of the homeward flight are predictable from those of the outward flight or perhaps because the dance codes, after all, an outward flight.

A method is outlined for determining the linguistic encoding of distance, a method primarily involving the application of a rigid structural and psychological simplicity criterion. The surviving energy drain correlations are to be examined to determine which are biologically dependent on each other; the most highly correlating element of each set of dependent energy drain measures is then hypothesized to be part of the linguistic encoding of distance. Neither the model bee nor the statistical measure of recruiting accuracy plays a role in this determination, although the question of the mode of self-monitoring of the dance is significant in this regard, as is the question of which correlation(s) are perceptible to the dance attenders. The von Frisch conclusion that the linguistic encoding of distance lies in the duration of wagging

(modified here to include the duration of sound production) may well ultimately be the preferred hypothesis but it remains to be shown that they are each their own set's best correlate with the energy drain, rather than that they are the overall best correlates with the map distance.

When the lighting is either diffuse or lacking, the azimuth of the goal is given with respect to the upward direction, provided the dance floor is sufficiently vertical to permit the dancer's gravity-sensing organs to function properly. The azimuth of the goal is given with respect to the azimuth of the sun whenever it is directly visible; when its position must be perceptually calculated from the sunlight between the clouds, the polarization pattern of the blue sky, a weighted average of the solar and upward axes is employed as axis, the weighting a product of the angle of inclination of the dance floor and the thickness of the cloud cover: a fully vertical dance floor provokes an even averaging of the by-light and by-gravity directions if the sky is visible. The translation from direction with respect to the sun to direction with respect to gravity is finally subject to a slight but systematic deflection toward the horizontal and vertical axes.

The positioning of potential dance recruits on the dance floor displays a diurnal rhythm, according to which the usual seekers of a particular product at a given time of day are at that time huddled together on a part of

the dance floor roughly corresponding to the location of their goal; this facilitates their alerting by a scout already engaged in harvesting that product. Successful dance following also involves the employment of a strategy; an angle of  $40^\circ$  is maintained between the dancer's and attender's bodies in such a fashion that the dance follower is alternately on the dancer's right and on her left, but consistently on the inside of the circuit.

Unusually long site advertisements (marathon dances) take account of the changing solar azimuth, the only aspect of the dance language known to be learned. The constituency of message repetitions is a serious problem, as is that of the round dance, whose reported "slovenliness" appears to be greater than that of the other dance types, although the vast array of completion curves in the S-dance is also problematic.

Although the dance language has been shown to be communicative, the discovery that recruiting is often a product of olfactory and visual cueing makes the interpretation of much of the literature, concerned in most cases with recruit behavior, difficult to interpret, and impossible to rely on. Many of the classic experiments need to be repeated within the misdirection paradigm of Gould 1975b (in which regard, see sec. 2.13).

It is likely that a large part of the account of dance production and comprehension will prove to be biologically

interesting but linguistically irrelevant. To take a case in point, the preference for the sunbound direction in the 180°-discrepancy trained transport experiments, regardless of whether that factor resulted from the pre- or post-feeding leg of the round trip, is likely to prove to be a consequence of the particular biological mediation of directional coding rather than a crucial bit of evidence for grammar-building or evaluation. Other biologically interesting, likely to be linguistically irrelevant, factors include the specification of the properties in a goal necessary (for a given hive condition) for the release of dancing; the possible dependence of distance correlation maintenance on the nature of the environmental factor responsible for the disoriented dance; the reception of the model bee.

Of particular interest linguistically is the determination of whether perceptual data other than her fatigue (as, for example, her visual estimate of her progress) play a role in determining the distance coded in the dance, thus supporting the distance rather than energy drain view of this parameter; whether the dialect preference for one mode of transitional dance over another is absolute or merely statistical, which would support the view that the dance language is a single system whose "dialects" are reflections of interracial biological differences.

The notion of dance competence is clear at this time only in its broadest outlines. There appear to be two racial dialects, differing in the structural patterns constituting the behavioral repertoire and so, requiring somewhat different grammatical treatments. Although it is the wagging run which generally provides the distance and direction information, the patterning of the dance involves other danced maneuvers whose distribution, at least, must also be accounted for. Each dance type is restricted in use to a particular portion of the flight range (it is conceivable that these portions overlap to some extent), from the round dance used for the shortest distances to the S-dance reserved for distances near the foraging limit.

## Notes

## 2

## PRELIMINARY OBSERVATIONS: FOCUS ON PERFORMANCE

1. The issues, none of which arise in quite the same sense for the zoologist, are discussed in detail in the rest of this section; this argumentation may be skipped, sec. 3 (The Finite State Account) read at the conclusion of this sketch (p. 17) provided the positions sketched here are clear, deemed reasonable, and the methodological issues uninteresting.

2. Arranged in order of appropriateness to gradually larger distance values. Thus, the round dance is used for goals very near (how near is "very near" depends on the particular racial strain) the hive or swarm cluster, the S-dance for those near the foraging limit (again, racially variable). The ~~wavy~~ indicates concurrent wagging.

3. The distance vs. flight range issue sketched earlier (p. 7) is further explained below (see pp. 41-43).

4. This is precisely the topic addressed in sec. 4.3. The experimental evidence that the dance language is in fact a communication system (see sec. 2.13) plays a role there in suggesting that the dance behavior is self-monitored and so has a cognitive base.

5. Von Frisch 1967a:243, 250-51, 266-67 and Gould 1975b:16 suggest that the likelihood of dancing may correlate inversely with the time taken up in the unburdening process. A returning forager's first task consists in finding "takers" for the disbursement of the load acquired at the feeding place.

6. The scout bee househunting during or just previous to swarming does not require separate treatment here. The major known differences between the forager's dance behavior and that of the scout bee are that the forager is in a position to offer food/water etc. samples to the dance attenders and that the scout bee on the average dances longer upon her return (this is discussed on p. 69 below).

7. It is difficult to know quite how to interpret the following quote, given the nature of the dance language controversy (see sec. 2.13): "Esch (1963) determined that in 6000 cases of silent wagging runs no newcomer came to the feeding place (at a distance of 150m). On the other hand it happened also that on rare occasions vibrating episodes occurred throughout the entire dance, even during the return phase. Such dances too were without success in recruiting" (von Frisch 1967a:58). Such dances are also apparently "feeble."

8. For a discussion of the progress made toward that end and an analysis of the problems still remaining, see Gould 1975b:62-75. In a similar vein, it has often been argued (see Limber 1976, for example) that the child's competence is not fully reflected in production data alone--comprehension data must be sought as well.

9. And, presumably, an analogous number-of-sound bursts correlation, although this parameter is not discussed.

10. Kroeber 1952 argues that the relevant parameter is developmental in nature: since the dance language's semanticity is genetically encoded, it is not arbitrary and so, not symbolic. That is not what is being addressed at this time. Rather, the question is that of determining whether there is a cognitive substrate to the duration of the wagging phase. An affirmative answer is suggested by a semantic interpretation treating that phase as coding distance, a negative answer by one treating it as coding energy drain/sugar consumption.

11. See von Frisch 1967a:296-97, which further raises "the question whether, when Carniolan and Italian bees are crossbred, the behavior of performing sickle dances is linked in inheritance with the yellow body marking," arguing that "the linkage of the mode of dancing with the color of the body should be reinvestigated . . . with Italian-Carniolan hybrid colonies whose queens have been inseminated artificially, so that the nature of the cross is absolutely clear," noting the complexity of the genetic coding of coloration.

12. One, that is, modelled on the misdirection experiments of Gould 1975b, outlined in sec. 2.13.

13. If the bee in question were sufficiently familiar with the area she was being transported to, the training period was quite brief or even unnecessary (von Frisch 1967a: 118).

14. The result could conceivably depend on whether or not the site of the "return" flight had previously functioned as a feeding place for the dancer in question.

15. With regard to reversal, witness the above-mentioned tendency to reverse the pre-feeding flight to return to the hive. With regard to averaging, see Gould 1975b: 170-73 for statistical comparisons of recruit distribution and production scatter, suggesting that the six or so consecutive pattern traversals witnessed by the successful recruit have been averaged. Also see p. 65 below.

16. Notably, the observed weighting of the two flights according to their length (the direction on the longer flight is more heavily weighted) and the weighting contributed by a line of landmarks along one of the hive routes (Otto 1959, reported in von Frisch 1967a:171).

17. Subject to the caveat described in the previous paragraph, it should be noted that even when the return flight is so heavily weighted by landmarks (see n. 16) as to be fully determinative of the dance-coded signal, the dancer's next foraging flight is modelled on the previous outbound flight (von Frisch 1967a:172n).

18. The direction-indicating segments of the dance are identified as the sites with wagging potential (see (1), p. 22).

19. "It turned out that the bee can orient just as well with regard to the polarized light of the blue sky whose direction of vibration depends on the position of the sun, as with respect to the sun itself" (von Frisch 1967a: 134).

20. Leading the forager to the feeding station by way of a detour around an object (building, hill, etc.) may result in her treating something as an obstacle which is not so perceived by the recruits, who keep to the air line, perhaps flying over rather than detouring around the "obstacle" (von Frisch 1967a:178). If the training is along a zig-zagging path without there being any candidate for the position of obstacle, the dancer eventually uses the air line on her outward as well as return flight; this mirrors the behavior of a bee who after an extended, irregular exploratory flight finally locates a profitable food source, returns to the hive along the air line, without retracing her inefficient outbound flight, her subsequent danced direction also pointing directly to the goal (see von Frisch 1967a:182-83).

21. In at least one respect this is an oversimplification. The provision of nectar and pollen naturally displays a diurnal rhythm which the bees are sensitive to. The same bee may be a member of as many as four foraging groups operating to harvest from four different locations active at four different times of the day (Finke 1958, reported in von Frisch 1967a:254). The dance floor then also displays a diurnal compositional rhythm in that any given foraging group assembles at the appropriate part of the dance floor at the proper hour, its members otherwise engaged (resting, etc.) the rest of the day.

22. Gould 1975a:690 (also see Gould 1975b:152-66) demonstrates that the statistical distribution of recruits argues against "the integration of separate dances," by which is meant the averaging of distance and directional indications produced by different dancers advertising different, olfactorily identical sites.

23. "Observations by Lindauer suggest that the shifting azimuthal angular velocity is not included in their calculations of the sun's nocturnal path" (von Frisch 1967a:365). It is not known how much of a role this plays even during daylight dancing.


24. This tendency toward error (roughly twice as strong a tendency toward vertical as toward horizontal directional deflection) is thus demonstrated for two toward-the-goal tasks. A correspondence between residual misdirection and a perturbation of a goal to hive angle (rather than, as here, a hive to goal angle) would have been difficult if not impossible to explain.

25. As by Gould's "misdirection" experimental paradigm--see discussion, sec. 2.13.

26. Correctly interpreted, that is, within the limits imposed by the 40° dancer-observer discrepancy discussed above.

27. The situation may be clarified by reference to the example (see "von Frisch's account") in (8a) below.

28. See (8b) below ("von Frisch's account") for clarification.

29. It is unlikely, though not beyond the realm of possibility, for the observer to correct for the  - estimate discrepancy produced by the approximately 40° orientation difference between the dancer and herself. (Note that the dancer cannot make this correction as the dance language is typically many observers-one dancer, with each

observer oriented differently to the dancer.) This would require something on the order of the lateral drift compensator discussed on p. 58. However, in the absence of experimental evidence suggesting tight agreement on the position of  $\odot$ , the hypothesis that the biological mechanism mediating the maintenance of the  $40^\circ$  angle also mediates a perceptual correction of  $\odot$  - estimates is not to be preferred to the account outlined here, due to Gould 1975b.

30. This very abbreviated outline of the dance language controversy is not intended to substitute for the historical analysis of Gould 1976, but only to sketch some of the major developments. As will be shown, the central issue with regard to the linguistic analysis of the dance language is not that of the system's communicative status, but rather its cognitive status. That is, the central issue is whether it is possible to motivate a competence/performance distinction, i.e., whether the dance language is, in Bennett 1964's terms, regular or rule-guided (see discussion, sec. 4.3).

31. Schricker 1974b shows that parathion also dissociates the "internal clock" of the poisoned forager.

### THE FINITE STATE APPROACH

As a competence model, the finite state device or Markov process (see, for example, McCulloch and Pitts 1943) characterizes the set of well-formed strings in the following way: any path from the designated initial state  $S_0$  to the designated final state  $S_f$  ( $S_0$  and  $S_f$  being elements of the finite set of states used as "way-stations" in the device) which agrees with the directionality of the arrows connecting the states in that path represents a well-formed string; nothing else is well-formed. The finite state device is required to generate the dance footwork rather than simply the resulting pattern, clearly distinguishes continuous dependencies from discontinuous ones. Consider, for example, the two wagging runs of the transitional or tailwagging dance; any dependency here would be discontinuous inasmuch as these two runs are necessarily separated in execution by a roughly semicircular run.

#### 3.1 Dances Indicating Distance and Direction: The Carniola Honey Bee

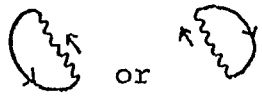
##### 3.1.1 The First Constituent of the Dance

The first problem that arises, therefore, is that of deciding which component of the dance is to be generated first. Based on von Frisch's (1967a:102) observation that

dance attenders who lose contact with the dancing forager or scout (see (6)) often wait for her where they anticipate the beginning of a wagging run, it is the wagging run which will be considered the first component of these dances.

While this appears inconclusive and perhaps even irrelevant at first, it is in fact inarguable on the assumption<sup>1</sup> that long intervals of dancing (to be referred to as superdances) are series of dances, where the relation between dance and superdance parallels that between sentence and paragraph or discourse monologue in human language, i.e., where it is dance and not superdance which the syntax must account for. Given this assumption, the dance follower who finds herself "out of the swing of things" and reacts by positioning herself so as to be in on a wagging run can only be interpreted as displaying her recognition that dances start with wagging runs.

This behavior will not represent the recognition of a constituent boundary D (dance) unless it can be demonstrated that a bee distracted on the first curved return run of a transitional or tailwagging dance does not join in the next wagging run but skips one. This seems highly unlikely in view of the fact that a bee may become an observer at any stage in the derivation of the SD (superdance) and so be unaware of which circuit (von Frisch's term for a wagging run together with its return, i.e.,

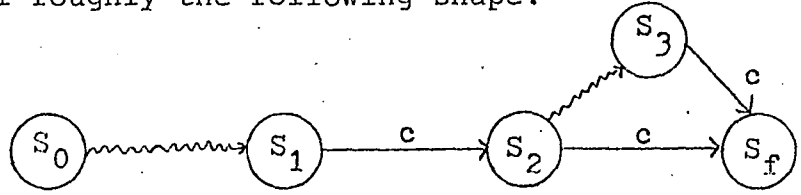


or )--left or right--the dancer has, so to speak, intended to be the first half of D.<sup>2</sup> The general point, however, that distance and direction indicating dances begin with a waggling run stands, provided only that the repositioning behavior described for the suddenly detached would-be dance attender holds across the various distance and direction indicating dance types (tailwagging:



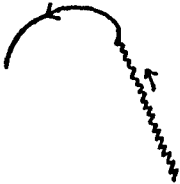

transitional, tailwagging, S-dance).


These observations serve to indicate a finite state device of roughly the following shape:





where  $S_0$  is the initial state,  $S_f$  the final state, denotes a waggling run and c a curved run. The path through  $S_3$  is intended to produce tailwagging transitional and (where the same waggling run is chosen in proceeding from  $S_2$  to  $S_3$  as from  $S_0$  to  $S_1$ ) tailwagging dances, while that going directly from  $S_2$  to  $S_f$  is to produce S-dances. Now note that this automaton, in not distinguishing one waggling run from another, fails to derivationally distinguish the tailwagging transitional and tailwagging dances. What's more, this automaton makes the claim that any waggling plus curved run combination is a viable beginning for any of

these dance types.




Yet  can only be completed as an S-dance and 

cannot be completed with a curve (\*  ). Thus there are dependencies between the various constituents in the dance, dependencies which the finite state device given above can't handle.

There are other criticisms one might level at this device, perhaps most importantly that in not distinguishing waggling runs of different durations or at different angles with respect to the axis (the axis being a ray directed toward the sun's azimuth or away from the earth, depending upon the dance's physical setting) it fails to make distinctions which will be important at the semantic (here necessarily equated with the information) level. The duration parameter additionally has a direct well-formedness correlate in that the two waggling runs (each of which is to give the distance or, more precisely, energy drain) must

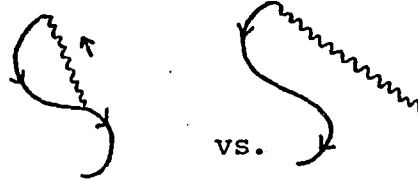
have the same duration, i.e. \*  , \* 

An adequate treatment must distinguish (a) signs differing along parameters which make semantic distinctions, e.g.


 vs.  and  vs.  , and (b) signs differing

along parameters which give rise to syntactic ill-formedness, e.g., curved runs must be specified so as to make it determinate whether  $c$  is a return to the place the last

wagging run began, as witness \*



As the illustrative examples show, wagging runs must be specified as to duration and angle, curved runs as to their relationship with the previous wagging run (is it a return run?) and finally (a type (b) example), in order to rule

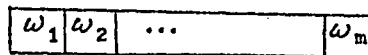
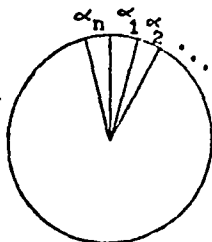
out \*  as a complete dance, the side of the wagging run the curve is run on.

### 3.1.2 The Lexicon

The following notation is accordingly employed:

$\omega_i \alpha_j$  denotes a wagging run of duration  $\omega_i$ ,  $\omega_i$  being one of a finite number of equivalence classes  $\omega_1, \omega_2, \dots, \omega_m$ , danced at an angle  $\alpha_j$ ,  $\alpha_j$  being one of a finite number of equivalence classes  $\alpha_1, \alpha_2, \dots, \alpha_n$ . The situation may be pictured as:

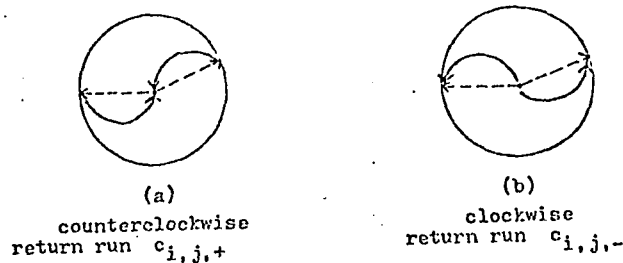
(9)



The two grids may be anchored at  $\omega_1 \alpha_1$ , the wagging run indicating the smallest danced distance value with the goal directly in line with the solar azimuth.


Curved runs will be described along three dimensions; the basic idea is that any curved run is a return run on a particular side of a wagging run which may or may not have been independently executed in that dance. In each of the following diagrams, the curve is interpreted as a return run to the broken ( - - - ) ray  $\omega_1 \alpha_1$  running from the center of the circle to the periphery; in (10a), the curve is drawn counterclockwise and written  $c_{i,j,+}$ , while in (10b) the curve is drawn clockwise and written  $c_{i,j,-}$  as indicated. It must be emphasized that the wagging run that


(10)

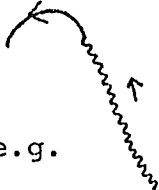



is associated with a given curve connects the end of the curve with its start, so that the wagging run for  $\curvearrowright$  and  $\curvearrowleft$  runs right to left while that for  $\curvearrowleft$  and  $\curvearrowright$  runs left to right. Attention may now be given to just which sequence of these are well-formed.

As discussed earlier in this section, the first component of the dance is a wagging run. The problem the finite state device must solve then is the specification of

what may follow a particular  $\omega_i \propto j$ . It is important in this regard to note that those cases where the curved run following  $\omega_i \propto j$  is a return run (e.g. ) can not

be completed with a simple curve (\* ) , while those

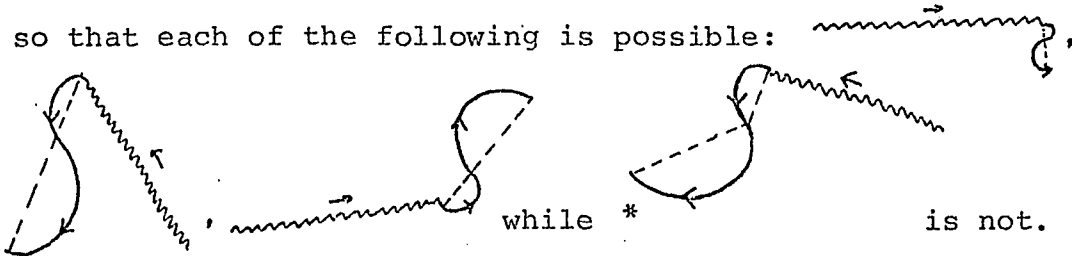
where the curved run is other than a return run (e.g. )

can only be so completed ( , \* .) This means that there will have to be different states  $W_{i,j}$  for each of the initial  $\omega_i \propto j$  runs so that in every case, depending upon the relationship the given curved run bears to  $\omega_i \propto j$  the appropriate completion sequence(s) (wagging run plus curve or simple curve) may be specified.

Now since the clockwise/counterclockwise specification of the curved run plays a rôle only in terms of ensuring opposite specifications for the two curved runs in a dance, attention will be further restricted to cases where the first curved run is counterclockwise.<sup>3</sup> The problem then consists in generating the proper completion sequences for a given  $\omega_i \propto j$  followed by a particular  $c_{p,q,+}$ .<sup>4</sup>

### 3.1.3 S-Dance

As discussed above, unless  $p=i$  and  $q=j$ , the sequence can be completed only by a simple curve, i.e. the dance must be an S-dance. Von Frisch's (1967a:59) discussion of this dance pattern, though brief, suggests that the size of the two curves need not be proportionate but that they must produce a concavity change about an axis, so that each of the following is possible:



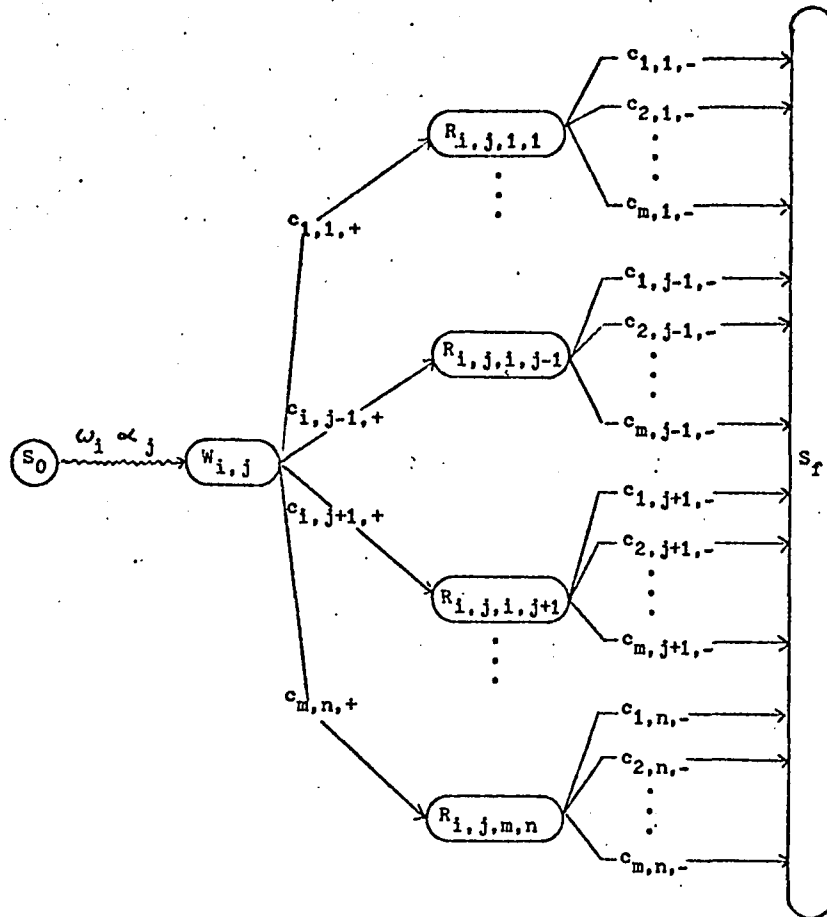
These conditions may be represented as follows in (11).

In each case the waggling run plus non-return counterclockwise run may be completed by any curved clockwise run made at the same angle as the first. In each case an S-dance is obtained.

### 3.1.4 Tailwagging Dance

When, on the other hand, a counterclockwise return run succeeds the initial waggling run, a second waggling run must be instituted. If it is identical to the first, a tailwagging dance results. If it is not, the dance is a transitional one. In either event, the second waggling run must be succeeded by a clockwise return. So, for the tailwagging dance, the device is as sketched in (12).

(11)



(12)




### 3.1.5 Transitional Dances: Y-Shaped

The transitional dance presents one other slight difficulty: no indication is given to suggest that



is grammatical. Assuming it is not, it must

be ensured that the wagging run to succeed  may have an angle component at most a fixed distance round (clockwise -- see (9)) from that of the first wagging run. Indeed, Gould 1976:216 reports that "some runs are wrong by as much as  $30^\circ$ ," confining the angular spread to  $60^\circ$ ,  $1/6$  of the angular range encompassed by the  $\alpha$ -variable. This observation may be taken as evidence that  $n$  (see (9)) is a multiple of 6.

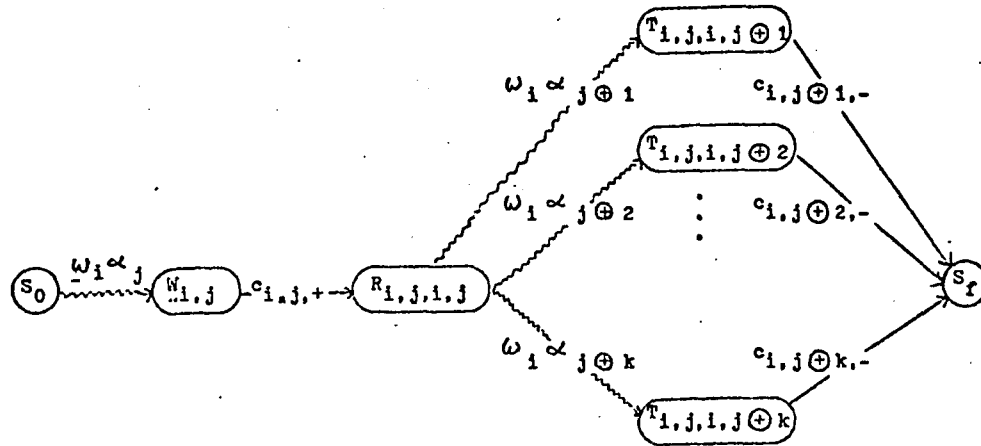
If  $\oplus$  denotes "clock addition,"<sup>5</sup> with for example  $(n-3) \oplus 4 = 1$ , the way ( $n=12$  on a clock) four hours past nine is one o'clock, then the limit can be set precisely using (13)'s  $k$  as the maximum positive deviation from  $j$

$$(13) \quad k = n/6$$

allowed the angular component of the second wagging run. Since  $n$  is a multiple of 6,  $k$  must be an integer. Tailwagging transitional dances can then be specified as in (14).<sup>6</sup>

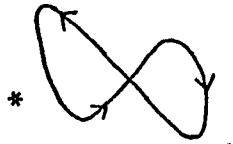
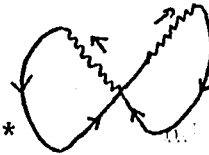
The incorporation of the remaining transitional dances of the Carniolan honey bee into this model crucially involves the representation of the straight run. Although straight-run duration has no semantic correlate, it is

(14)



necessary to make the same distinctions as for the wagging

run in order to rule out such sequences as \*



In terms of the model already given and the notation  $S_{i,j}$  for a straight run of the same direction and duration as  $\omega_i \alpha_j$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n$ , such dances can be generated by introducing a new exit to the state  $R_{i,j,r,s}$  (see (11)), permitting the generation of such sequences as (15).

(15)

$$\omega_i \alpha_j \quad c_{r,s,+} \quad S_{p,q} \quad \omega_i \alpha_q \quad c_{r,u,-} \quad S_{p,j}$$

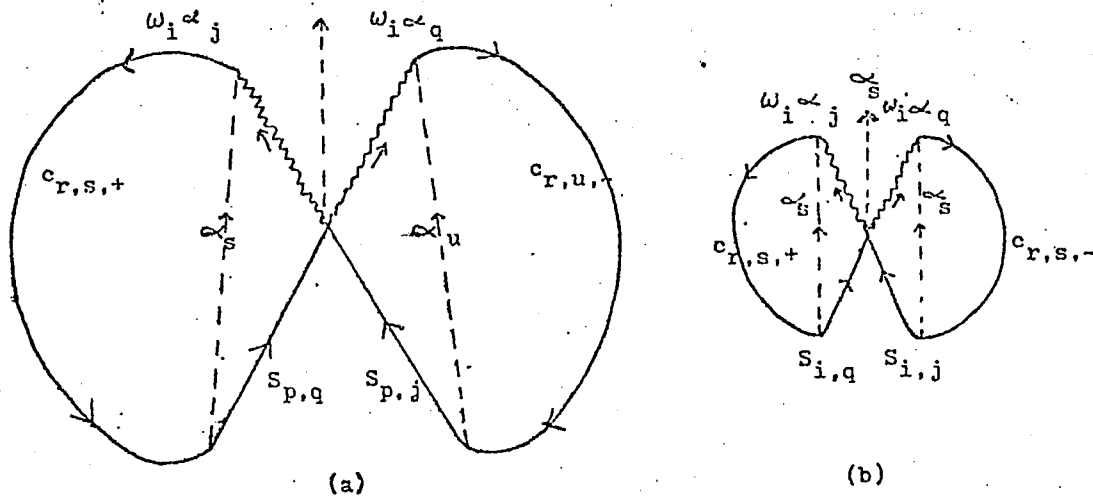
In each case, the following conditions will obtain:

$s \neq j$ ,  $\alpha_q$  is less than  $60^\circ$  ( $k$  units) away from  $\alpha_j$ ,  $S_{p,q}$

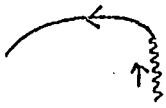
completes the circuit (returning the dancer to the starting

point of her initial wagging run ( $\omega_i \propto_j$ ), and finally,  $c_{r,u,-}$  is that curved run which will allow  $S_{p,j}$  to complete the second circuit (again returning the dancer to the starting point of her initial  $\omega_i \propto_j$ ), as in (16a). When  $p = i$ ,  $\propto_s$  corresponds to the bisector of the two wagging runs,  $\omega_i \propto_j$  and  $\omega_i \propto_q$ , and  $u = s$ . More importantly, such a case (see (16b)) with  $q = j \oplus k$  would show  $k$  to be a multiple of 2 and thus  $n$  (see (13)) to be a multiple of 12.

(16)



The  $60^\circ$  condition acts to preclude straight-run

transitional dance completion of  or any other sequence whose transitional completion would result in the direction's being given by the bisector of wagging runs more than  $60^\circ$  apart. We assume that Gould's observation

(see p. 108) holds across dance types and dialects. Ideally, the 60° condition should be given in terms of  $\alpha_s$ , identifying which  $\omega_i \alpha_j c_{r,s,+}$  sequences are completable as transitional dances. However, this is in a very real sense unimportant here. Indeed, the fact that the values for  $p, q, t, u$  are computable from  $i, j, r, s$  is in itself irrelevant, useful only for expository purposes. The uniquely satisfying  $p, q, t, u$  for a transitional  $i, j, r, s$  quadruple must be specified directly in the automaton rather than by a formula.

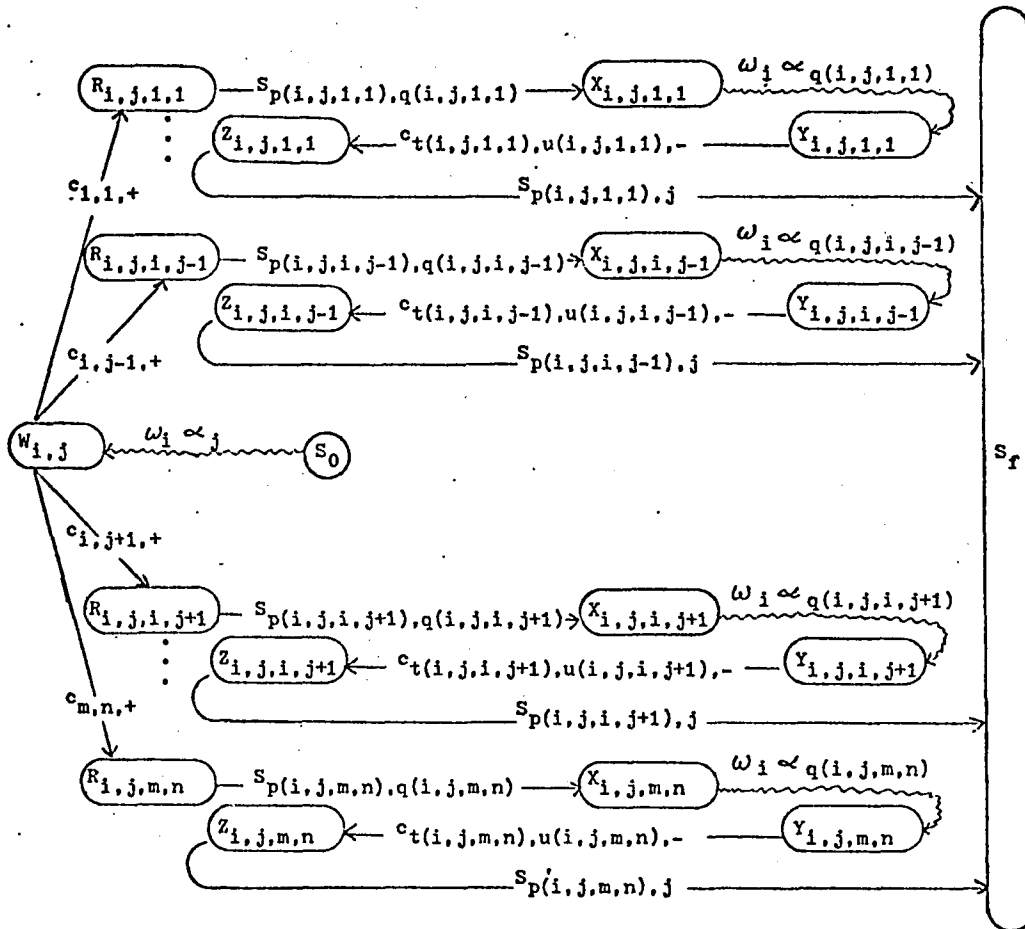
The finite state device is therefore to be enlarged along the lines sketched in (17), assuming that the paths indicated meet the circuit-completing and  $s \neq j$  conditions, and that those featuring  $q$  values other than  $j \oplus 1, j \oplus 2, \dots, j \oplus k$  have been eliminated. Where  $p(i, q, r, s) = i$ ,  $q \in \{j \oplus 2, j \oplus 4, \dots, j \oplus k\}$ , as discussed above with respect to (16b).

Non-wagging Y-shaped transitional dances may be accounted for by the addition of a device very like that in (17), save that  $S_{x,y}$  substitutes for  $\omega_x \alpha_y$  throughout, as in (18). The same restrictions, of course, apply here as for (17).

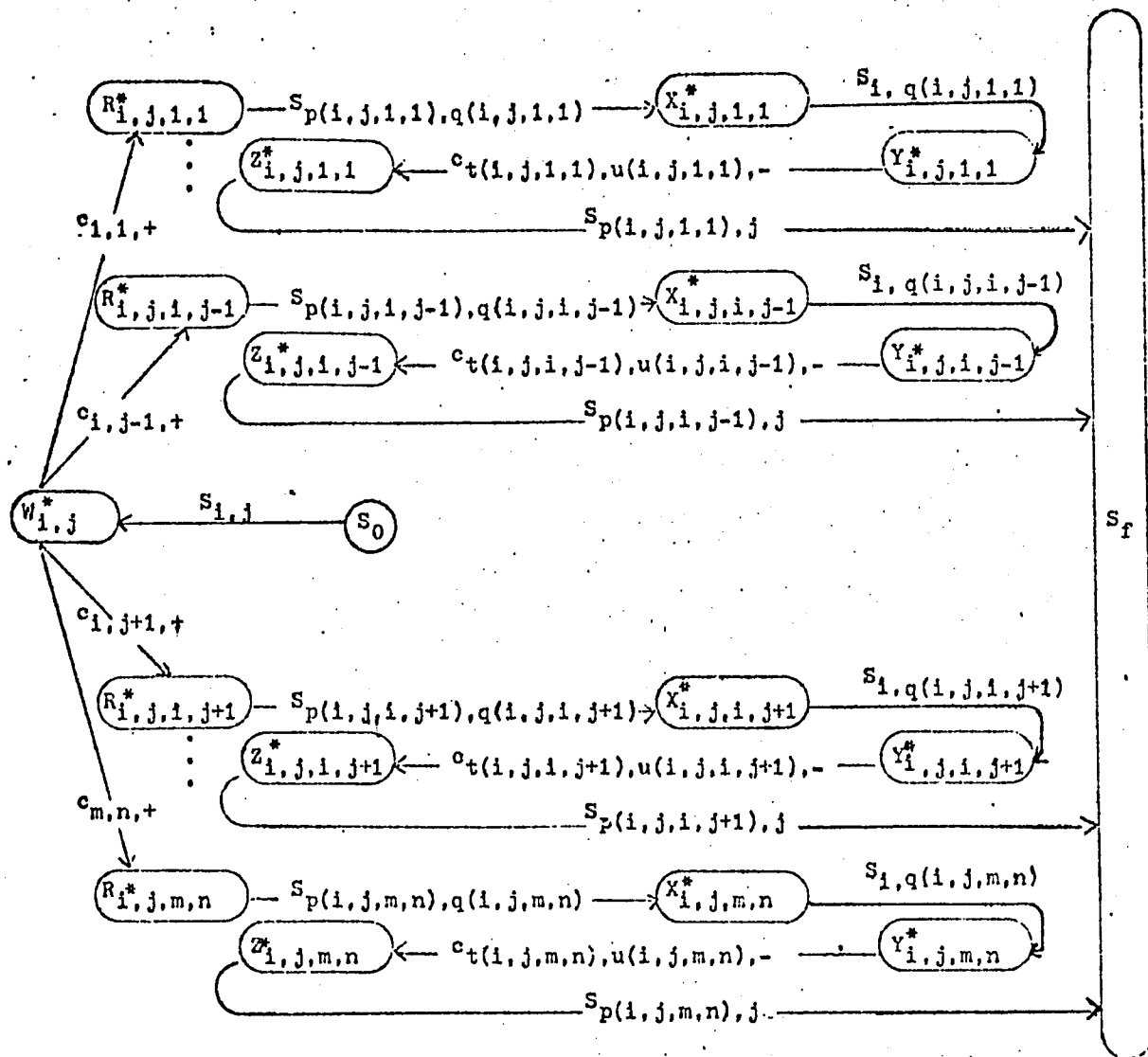
### 3.2 The Grammar Thus Far

The finite state account of the Carniolan honey bee's distance and/or direction indicating dances is thus an

(17)



(18)



amalgam of (11), (12), (14) and the  $s \neq j$ , circuit completing, k-modified (17) and (18), together with their mirror-image counterparts constructed as outlined in n.3.

Wagging phases (or in the case of the nonwagging transitional dance, sites suitable for wagging) are invariably generated as the initial dance component, on the strength of von Frisch's (1967a:102) report (see p. 100) that breaks in dance following are repaired where wagging runs begin, an observation which needs to be confirmed across dance types before this tactic is fully justified.

The finite state model must eventually be examined to eliminate paths combining a wagging/sound production duration value incompatible with the dance type the path is producing. As already discussed, each dance type accesses a fixed range of wagging/sound production duration values (see pp. 23-24, these ranges possibly overlapping (see pp. 67-68)).

Given that the semantic interpretation of transitional dances involves averaging the different directional indications, and that the  $\alpha$ -range of (9) represents a conceptual organization, it must further be guaranteed that the averaging produces a value within the  $\alpha$ -range. To take a concrete example, a tailwagging transitional dance of the form  $\omega_i \alpha_j c_{i,j,+} \omega_i \alpha_{j \oplus 1} c_{i,j \oplus 1,-}$  must be supposed impossible since it indicates a direction of  $j \oplus 1/2$ , not an element of the  $\alpha$ -range. This point may be put another

way: there is no principled reason for allowing transitional dances a greater directional accuracy than the tailwagging or S-dance. The k-modification clause for (17), (18) which requires that  $q \in \{j\oplus 1, j\oplus 2, \dots, j\oplus k\}$  is not sufficient;  $q$  must be an even number of units from  $j$ ; the bisector must be an element of  $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ . In other words, the k-modification clause should read  $q \in \{j\oplus 2, j\oplus 4, \dots, j\oplus k\}$ ,  $k$  even, whether or not  $p=i$ . Similarly, the exits from  $(R_{i,j,i,j})$  in (14) should be restricted to wagging runs an even number of units from  $j$ , i.e., to the values  $j\oplus 2, j\oplus 4, \dots, j\oplus k$ .

### 3.3 Dances Indicating Distance and Direction: The Italian Honey Bee

Assuming that von Frisch's observation with regard to the repair of breaks in dance following holds not only across dance types, but also across races of honey bee, the fact that the Italian race shares both the S- and tailwagging dance with the Carniolan honey bee is easily accommodated by attributing to the two models subparts of the sort presented in (11) and (12), together of course with their mirror-image counterparts (see n. 3). The correction for incompatible wagging/sound production duration values and dance types is liable to prove to be dialect-independent given the previous discussion of the relationship between these values and the percentage of the flight

range (see pp. 41-43).

Presumably, identical conclusions should be reached as to the status and structure of the round dance, discussed below. The unresolved distance vs. flight range percentage issue plays a crucial role here in determining whether the referential difference between the two dialects' view of a particular  $\omega_i \propto_j$  is semantic or not. If, as traditionally claimed, it is distance which is encoded, this difference is semantic, with a greater distance assigned  $\omega_i \propto_j$  by the Carniolan lexicon than by the Italian (see p. 23). Gould's  $60^\circ$  limit on directional variation, based on observation of Italian bees, is naturally retained here.

It thus remains only to consider how to generate the transitional dances typical of the Italian honey bee.

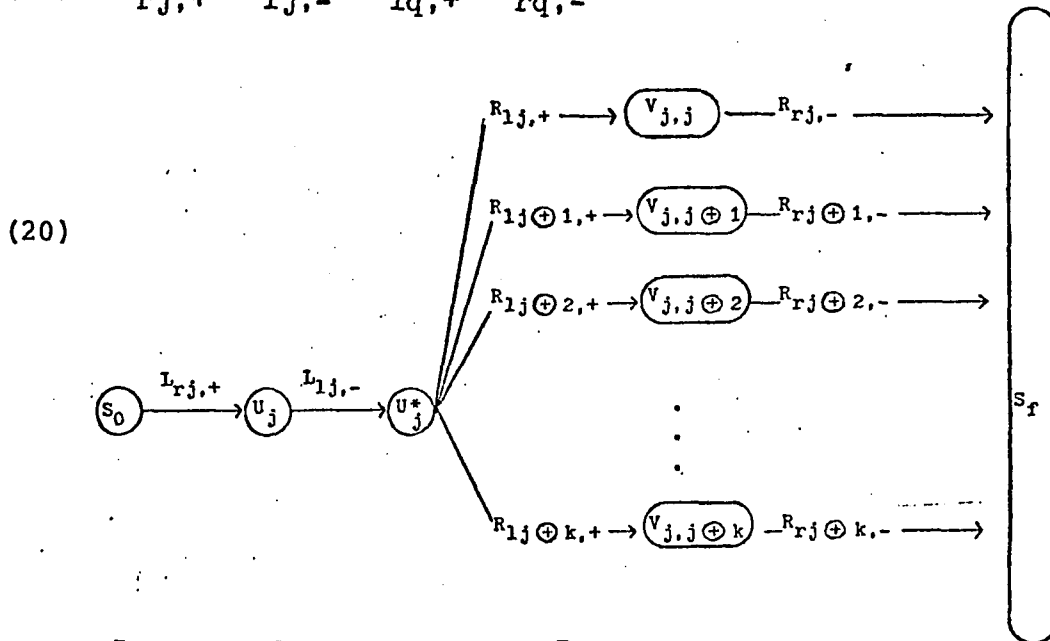
### 3.3.1 Transitional Dances: Sickle-Shaped

The sickle dance is characterized by four curved portions, none of which may be neatly paired as were the clockwise vs. counterclockwise semicircular return runs of the tailwagging dance. The vocabulary thus specifies four curved counterparts for a wagging run  $\omega_i \propto_j$ ; the left side of the left branch of a sickle based on  $\propto_j$  as  $L_{lj}$ , the right side of the left branch as  $L_{rj}$ , the left side of the right branch as  $R_{lj}$ , and the right branch of the right branch as  $R_{rj}$ .  $L_{rj}$  and  $R_{lj}$ , as the inside curves of the sickle, may be further specified as  $L_{ri,j}$  and  $R_{li,j}$  to

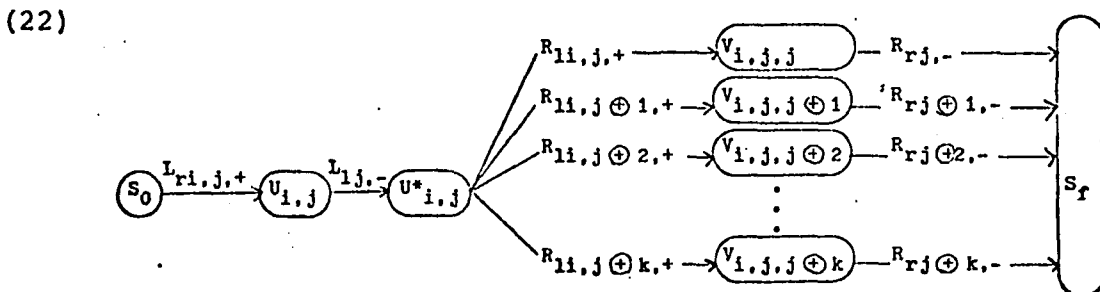
indicate concurrent wagging (at the open end of the formation) of duration  $\omega_i$ . Finally, each curve is specified as + if run away from the four-curve contact point, - if run toward that point.

The nonwagging sickle (19) is generated as in (20), the wagging sickle (21) as in (22).

(19)  $L_{rj,+} \quad L_{lj,-} \quad R_{lq,+} \quad R_{rq,-}$

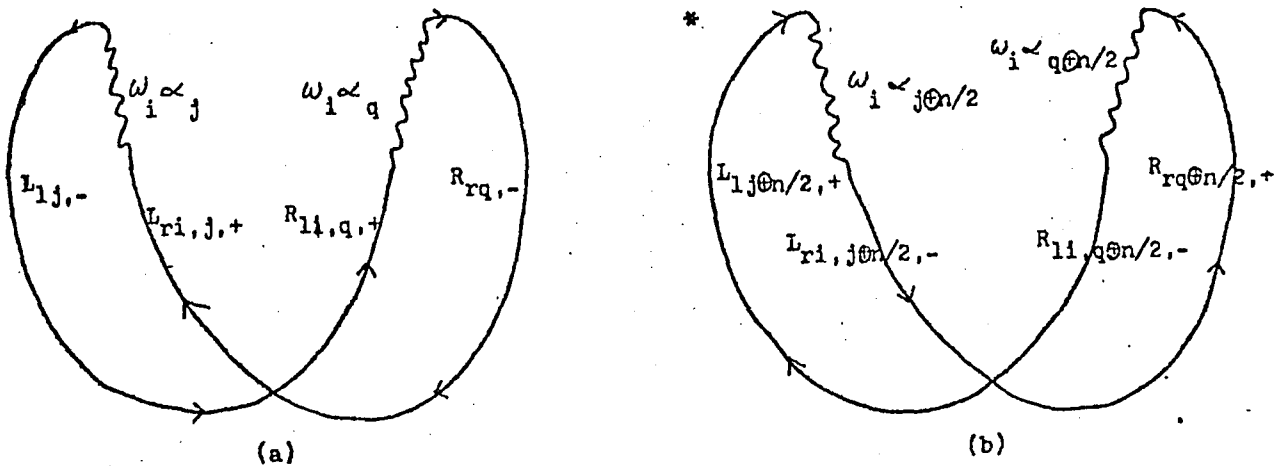


(21)  $L_{ri,j,+} \quad L_{lj,-} \quad R_{li,q,+} \quad R_{rq,-}$



Only the counterclockwise generation of the basic pattern is outlined here (see (23a)). The reverse traversal (23b) of this pattern is ungrammatical; the mirroring procedure of n.3 is blocked here as in (14) -- see n.6 for discussion.

(23)



It is unclear whether the sickle dance may take the full-horseshoe shape of (24a,b), but such dances are easily excluded should that prove necessary.

(24a)  $L_{rj,+}$   $L_{lj,-}$   $R_{lj,+}$   $R_{rj,-}$

(b)  $L_{ri,j,+}$   $L_{lj,-}$   $R_{li,j,+}$   $R_{rj,-}$

### 3.4 The Round Dance

There is very little which can be said of a specific nature with regard to the structure of the round dance. The notational vocabulary thus far introduced is not necessarily inadequate; it is the nature and extent of the idealization to be made, the quality of the competence "tidying" of

its reported slovenliness in performance which is at issue. Indeed, this slovenliness is even worse than has thus far (see p. 61) been reported here, for not only does it extend to the irregular spacing/timing of her reversals but also to a reported possibility of concurrent, again irregularly spaced, coincident wagging/noise bursts during circling (von Frisch 1967a:45, 238). These bursts are taken there to be paralinguistic, indicative of the unusually good quality of the goal, part of the dancer's "vivacity," which during other dance forms is probably measurable in part by the lateral extent of her wagging movements (see Gould 1976:240). Provided that these phases (which, since overlaid on a circular run, can not code the direction) are moreover of irregular duration (and thus do not code the distance), there will be no reason to dispute this conclusion.

Given the information-coding underpinnings of the current treatment, whatever features fail to reliably encode information or to pattern in a distinctive fashion (e.g. the curved runs) are thereby not to be treated as linguistic, i.e. are not part of the dance proper. This in turn makes the treatment of the so-called direction-indicating round dance of Apis indica a matter of some moment. In this species, the reversal sites roughly approximate the direction to the goal (von Frisch 1967a:304). It is unclear from that discussion whether the directional

imprecision here is on the order of that occurring across sickle dances, or even whether non-direction-indicating round dances ever occur. It is distinctly possible that the "direction-indicating round dance" is nothing more than a superdance interpolating sickle dances among the round dances. The fact that such superdances are reported for goals as little as 10 cm. from the hive merely reflects this species' very small flight range. The energy drain range appropriated by this species for the round dance may be so small as to make use of the round dance very limited indeed.

Treating the round dance pattern as a full circle allows the irregularity in reversal to be treated as either an adjacency phenomenon involving deletion or as due to the interference of psycho-/physiological factors with dance execution.<sup>7</sup> There is some slight confirmation for this view in that nothing in the literature suggests that a superdance which does not fairly rapidly present a full-circle pattern (e.g., one consisting of rapid reversals on the order of \*  $\left( \uparrow \downarrow \uparrow \downarrow \dots \right)$  would be a possibility.

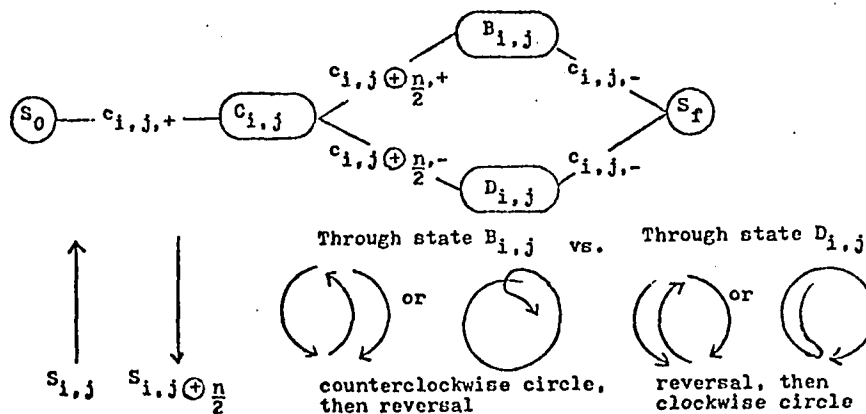
This construal of the status and structure of the round dance is easily accommodated. The curve  $c_{i,j,=}$ , by definition a counterclockwise return run to a non-executed  $w_i \alpha_j$ , is completed by a counterclockwise return run  $c_{i,h,+}$  drawn to a non-executed  $w_i \alpha_h$ , a waggling run identical to  $w_i \alpha_j$  save for having the opposite orientation. Again

supposing that no limitation on the regularity of the pattern is imposed by the vocabulary's equivalence classes (see (9)),  $h$  may be given as in (25) and  $n$  is presumed even independently of the specification of  $k$  (see (13)).

(25) 
$$h = j \oplus (\frac{n}{2})$$

The emphasis on reversal possibilities, however, further suggests that superdances consisting of a series of all-counterclockwise (or all-clockwise) tracings of the pattern might not occur. Further supposing such to be ungrammatical, not due to psycho-/physiological factors, suggests that the basic pattern be taken to include a full circle and a single reversal, this reversal either preceding or following the execution of the full circle, as in (26). The frequency of reversals in a superdance,

(26)



like their irregularity, is to be attributed, as per n.1, to performance factors. Again n.3 serves to produce the oppositely oriented cases (viz., clockwise circle plus reversal and reversal plus counterclockwise circle).

Further speculation is fruitless, although it must be noted that the best account of this dance type may require new vocabulary, specifying the round dance as a sequence, not of half-circles but rather of third- (or quarter- etc.) circles. Neither is it the case that plausible alternative analyses are available only for this dance type.

Consider the non-wagging Y-shaped transitional dance analyzed in (18), by analogy with the wagging Y-shaped transitional dance analyzed in (17), as a sequence of six segments, four straight and two curved runs. The straight runs  $S_{i,j}$   $S_{p,j}$  could quite conceivably be better described as a single straight run  $S_{i+p,j}$ ;  $S_{i,q}$  and  $S_{p,q}$  as  $S_{i+p,q}$ . What is needed is evidence having first of all to do with where dance followers who lose contact with the dancer await her (if at the crossing of the straight runs, this would argue for a four-straight-runs analysis) and secondly, whether  $i + p$  must be less than or equal to  $m$  (which would argue for a two-straight-runs analysis). Indeed, this dance type may not involve straight runs at all, but consist solely of curved runs, resembling nothing so much as a figure-eight when properly idealized. The

analysis then would be along the lines of (27) and (28),

$$(27) \quad c_{i,j,+} \quad c_{i,j \oplus \frac{n}{2},+} \quad c_{i,j \oplus \frac{n}{2},-} \quad c_{i,j,-}$$

$$(28) \quad c_{i,j,+} \quad c_{i,j,-} \quad c_{1,j \oplus \frac{n}{2},-} \quad c_{i,j \oplus \frac{n}{2},+}$$

depending upon which production order(s) occur. Perhaps there are no straight runs in any of the dances. It simply is not clear from the data available.

Finally, the discovery that an adequate analysis of the round dance is in terms of quarter- (or sixth-, or eighth-, ...) circles would suggest a reanalysis of the other dance types in those terms, while one in terms of third- (or fifth-, or seventh-, ...) circles would prove quite puzzling. Nor is it obvious what would constitute evidence in favor of any of these idealizations of the round dance over any of the others. The analysis proposed in (26) has the merit of being simple while incorporating two relatively reliable structural claims: an obligatory reversal and a full circle.

## Notes

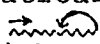

3

## THE FINITE STATE ACCOUNT

1. The idealizations hypothesized in constructing the finite state device below will be examined in secs. 4.1-4.4.

2. In the event that this is predictable (whether for a species, hive, or individual, or on the basis of some semantic content) this last argument does not, of course, go through.




3. The mirror-image cases will be handled by the ultimate addition of a state  $S'$  for every state  $S$  to be discussed. For any two states  $S_a, S_b$  in the restricted finite state device for which there is a path from  $S_a$  to  $S_b$  generating a counterclockwise movement (such as  $c_{i,j,+}$ , for example), there will be a path generating the motion clockwise ( $c_{i,j,-}$ , following the example) from  $S_a'$  to  $S_b'$ , and vice-versa. If there is a path from  $S_a$  to  $S_b$  generating any other sort of movement (such as  $\omega_i \alpha_j$ ), then there is such a path from  $S_a'$  to  $S_b'$ . Finally, if  $S_0$  and  $S_f$  are the designated initial and final states of the restricted finite state device, then  $S_0' = S_0$  and  $S_f' = S_f$ .


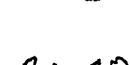
4. This assumes that every wagging plus curve sequence is completable, i.e. the initial part of some grammatical string. If in fact "short" or "long" return runs ( or ) must be disallowed (an empirical issue without, as far as I can see, any unfortunate consequences for what is to follow), the necessity of a different state for each of the  $mn$  distinct wagging runs beginning a dance is further supported.

5. This can be expressed in standard mathematical notation in the following way:


$$x \oplus y = \begin{cases} x +_n y, & \text{if } x +_n y > 0 \\ n, & \text{otherwise} \end{cases}$$

where  $+_n$  is addition modulo  $n$ , which may be defined as the remainder upon division of the sum by  $n$  ( $n > 0$ ).


6. Taking this 60° limit as indicative not only that  is ungrammatical (as on p. 108) but also that  is ungrammatical although it is the reverse traversal of the grammatical , requires that n.3 not be permitted to apply to (14). That is, the grammaticality of:

\*  $\omega_i^{\alpha_j} c_{i,j,-} \omega_i^{\alpha_{j\oplus 1}} c_{i,j\oplus 1,+}$  \*  

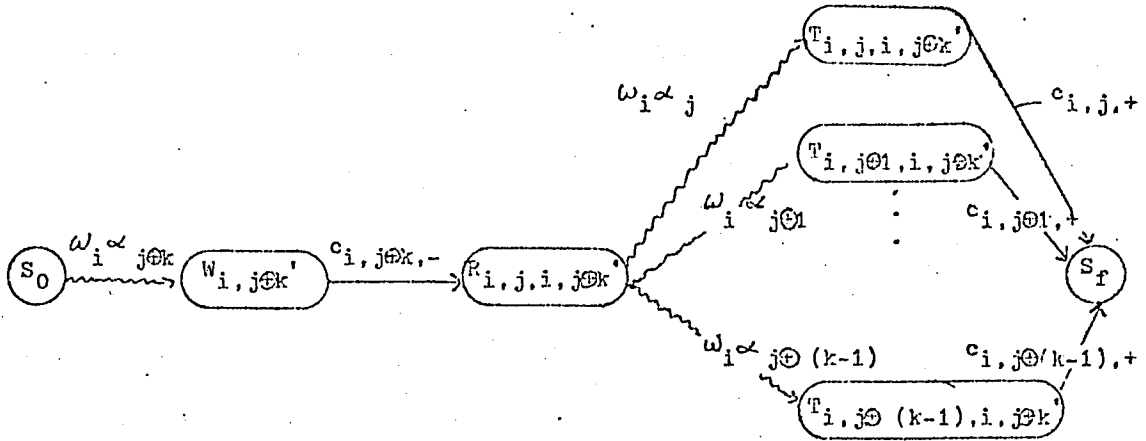
does not follow from that of:

$\omega_i^{\alpha_j} c_{i,j,+} \omega_i^{\alpha_{j\oplus 1}} c_{i,j\oplus 1,-}$  

What does follow, however, and moreover would not be characterized as grammatical by a device built analogously to (14), is:

$\omega_i^{\alpha_{j\oplus 1}} c_{i,j\oplus 1,-} \omega_i^{\alpha_j} c_{i,j,+}$  

the right-circuit first variant of the cited grammatical dance. (14) is accordingly augmented, not as per n. 3, but rather as follows:



This, rather than n.3, should be taken as the model for any other transitional dances with a 60° limit in directional spread.

7. On the face of it, this second explanation (in which a bee, reversing after, say,  $3/4$  of a revolution is seen as "losing her place") seems more plausible. The competence/performance question is further investigated in sec. 4.4.

## ALTERNATIVE GRAMMATICAL ACCOUNTS

Sample paths for the finite state automaton proposed for the Carniolan race are presented in (11), (12), (14), (26), and the modified (see pp. 109-111) versions of (17) and (18), for the Italian race in (11), (12), (14), (20), (22), and (26). Each device is then fully specified by giving the unaccounted for alternate traversals of the pattern as per sec. 3's n. 3 and n. 6, and finally, by giving the precise paths appropriate in the given racial dialect (see p. 23) for each particular value of  $i, j, p, r, \dots$ , obtained by excluding incompatible dance-type wagging- (sound-) duration pairs (so, excluding all but dialect-defined "far"  $i$  values for S-dances, and so forth).

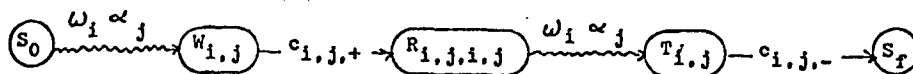
Each automaton corresponds to a grammar featuring rules of the form (29), where  $S_1$  and  $S_2$  are states of the automaton (such as, for example,  $S_f$  and  $W_{1,1}$ ),  $\rightarrow$  is read

$$(29) \quad S_1 \rightarrow a S_2$$

as "yields" or "is a" or "is rewritten as," and  $a$  is one of the smallest syntactic units of the dance (such as, for example,  $w_4 \alpha_6$  or  $c_{12,4,-}$ ). The sample tailwagging paths of (12), for example, repeated here as (30), correspond to

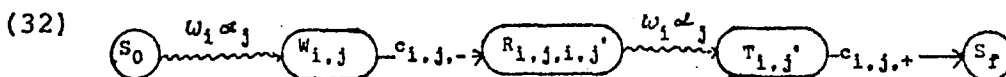
the series of rules in (31), ordered, as before, by specifying  $S_0$  as the initial (start) symbol, in effect the

(30)=(12)



$$(31) \quad \begin{aligned} S_0 &\rightarrow \omega_i \alpha_j W_{i,j} \\ W_{i,j} &\rightarrow c_{i,j,+} R_{i,j,i,j} \\ R_{i,j,i,j} &\rightarrow \omega_i \alpha_j T_{i,j} \\ T_{i,j} &\rightarrow c_{i,j,-} S_f \end{aligned}$$

symbol meaning "dance." Again, full specification depends upon rules of the form (31) for each value of  $i$  and  $j$  appropriate to the tailwagging dance (i.e., all values of  $j$  between 1 and  $n$ , and fairly large, but not too large, values of  $i$  between 1 and  $m$ ) both for this counter-clockwise circuit order and for (33), which additional rules devolve



$$(33) \quad \begin{aligned} S_0 &\rightarrow \omega_i \alpha_j W_{i,j}' \\ W_{i,j}' &\rightarrow c_{i,j,-} R_{i,j,i,j}' \\ R_{i,j,i,j}' &\rightarrow \omega_i \alpha_j T_{i,j}' \\ T_{i,j}' &\rightarrow c_{i,j,+} S_f \end{aligned}$$

from the mirror-image case added to (30)=(12) in (32), using sec. 3's n.3. Thus the sequence of runs composing a

grammatical dance are specified as those sequences of minimal syntactic units ("morphemes") generable from the dance symbol  $S_0$  using the provided rules. The tailwagging mini-grammar consisting of the rules in (31) and (33) will generate two tailwagging dances, derived in (34) and (35). Each line of (34) and (35) is derived from the previous one

$$\begin{aligned}
 (34) \quad & S_0 \\
 & \omega_i \alpha_j \quad W_{i,j} \\
 & \omega_i \alpha_j \quad c_{i,j,+} \quad R_{i,j,i,j} \\
 & \omega_i \alpha_j \quad c_{i,j,+} \quad \omega_i \alpha_j \quad T_{i,j} \\
 & \omega_i \alpha_j \quad c_{i,j,+} \quad \omega_i \alpha_j \quad c_{i,j,-} \quad S_f
 \end{aligned}$$

$$\begin{aligned}
 (35) \quad & S_0 \\
 & \omega_i \alpha_j \quad W_{i,j}' \\
 & \omega_i \alpha_j \quad c_{i,j,-} \quad R_{i,j,i,j}' \\
 & \omega_i \alpha_j \quad c_{i,j,-} \quad \omega_i \alpha_j \quad T_{i,j}' \\
 & \omega_i \alpha_j \quad c_{i,j,-} \quad \omega_i \alpha_j \quad c_{i,j,+} \quad S_f
 \end{aligned}$$

by the application of one of the rules in (31) and (33), respectively; the derivation terminates when the only category (state) symbol remaining is  $S_f$ : there is no rule of the form  $S_f \rightarrow \dots$ . Expanding (31) and (33) across the dialect-appropriate values of  $i$  and  $j$  will permit the generation of all grammatical tailwagging dances.

This labored description of the relationship between the finite state device and the grammar consisting of rules of the fixed type (29) is intended to justify the departure

at this time from the consideration of automata of varied design to the consideration of grammars of varied design (see Levelt 1974a: 63-4; for a full account of their equivalence). This departure is motivated solely on heuristic grounds, these being that only in this Markov process case is the design of the automata (given here as a transition diagram) intuitively more accessible than the design of the grammar.

The recognized grammatical alternatives<sup>1</sup> to the finite state schema of (29) are characterized with reference to the structural properties of the rules they feature; the resulting four-way distinction in (36) is due to Chomsky 1959. In what follows, A,B,... are category (state) symbols, a,b,... are morpheme symbols, and X,Y,... are finitely long strings of category and morpheme symbols.  $\lambda$  denotes the empty string, i.e., the string whose length is zero.

(36) Type 3. Finite State.  $A \rightarrow a B$

Type 2. Context Free Phrase Structure.  $A \rightarrow X, X \neq \lambda$

Type 1. Context Sensitive Phrase Structure.  $XAY \rightarrow XZY, Z \neq \lambda$

Type 0. Unrestricted Rewriting.  $X \rightarrow Y$

It has been shown (see Chomsky 1963 and Chomsky and Miller 1963) that each change in rule type changes the set of languages (see sec. 1, n.3) which can be characterized. That is, for any given alphabet, there are grammatical vs. ungrammatical bifurcations of the set of all finite concatenations

on that alphabet such that the grammatical subset (i.e., the language) is not generable using only rules of Type  $n$ , although Type  $n-1$  will generate it ( $1 \leq n \leq 3$ ).

Transformational grammar is the most widely accepted theory of human language, the need for transformations perhaps the most widely accepted example of a syntactic universal. The limitations of this theory as a psychological account of the idealized product of language acquisition (i.e., as a theory of competence) are clear, given the demonstration (Peters and Ritchie 1973) that transformational grammar is equivalent to an unrestricted rewriting system (Type 0)--in principle able to generate any recursively enumerable set (any set, that is, for which an effective characterizing procedure of any sort whatsoever exists) and not, as is desired, solely those recursively enumerable sets which have the property of being a notational variant of a possible human language.

Transformational grammar is the most widely accepted theory of human language, then, due to the demonstration that Types 1, 2, 3 are at best descriptively inadequate linguistic theories; in sec. 4.1 the arguments that Types 2, 3 are additionally observationally inadequate (see sec. 1 n.2 for the definition of levels of adequacy) are briefly sketched. It remains the major syntactic concern in contemporary linguistics to mathematically constrain the theory of transformational grammar so as to obtain a more

revealing characterization of human language. Numerous methodological constraints have been proposed (e.g., the notion of "independent evidence" for a hypothesized rule) only to be shown to be mathematically vacuous. Most approaches still agree that the syntactic component of the grammar consists of a series of rules of Type 2 followed by a series of rules ("transformations") of Type 0, which delete, move, insert lexical material in terms of their characterization (as noun or verb, for instance) as per the Type 2 rules (see, for example, Akmajian and Heny 1975).

The discussion thus far of the honey bee's dance language has served to provide a Type 3 account. The investigation now focusses upon the adequacy of that account, particularly in light of the competence/performance distinction which has proven so productive in the study of English syntax.

Any criticisms of the finite state device hypothesized will have to establish that (a) the account fails to capture structural information, fails to make generalizations that are linguistically significant or (b) the idealizations involved are not warranted, i.e., the data base is substantially different from that indicated. These two approaches to the evaluation of the finite state device are considered in the next two sections.

## 4.1 Structural Criticisms of the Finite State Account

### 4.1.1 Observational Adequacy

The work of sec. 3 rests on two crucial assumptions which together act to ensure that the dance language is of finite size and so, trivially amenable to a finite state treatment. The observational adequacy of the finite state approach is guaranteed provided that (a) there is a finite number of  $\alpha$  and  $\omega$  values and therefore a finite lexicon consisting of the possible  $\omega_i \alpha_j$ ,  $c_{i,j,+}$ ,  $s_{i,j}$ ,  $L_{ri,j}$  ... "morphemes"; and that (b) the syntactic unit (the dance) is the minimal message unit which features all the obligatory maneuvers<sup>2</sup> demonstrated in the message series (the superdance). The first of these assumptions is the topic of sec. 4.2; the second will be considered in some detail in this section. Primarily, however, this structural evaluation of the finite state approach must be in terms of meeting the goal of descriptive and not observational adequacy. Indeed, there is some reason to believe that the dance/superdance decision is immaterial to the determination of the power of the grammar, i.e. that even if the minimal syntactic unit were in fact the superdance, finite state capacity is (observationally) adequate.

The claim that it is superdance (SD) and not dance (D) which is to be generated may be interpreted as insisting on generating and generalizing from the entire message sequences produced in natural circumstances. It functions

to make the grammar directly responsible for handling what sec. 3 treated as adjacency phenomena. It is an open question whether the construction of this claim as  $SD \rightarrow DD^*$ <sup>3</sup> (a direct translation of adjacency questions from discourse analysis to syntax or, less formally, the treatment of SD as a conjunction of Ds) or the elimination of the constituent D is to be preferred. In the latter case, the semantic component would have to process any grammatical sequence in one blow; in the former case, Ds might be the input to the semantic component which could be viewed as having both D-interpretation rules and a conjunctions analyzer.

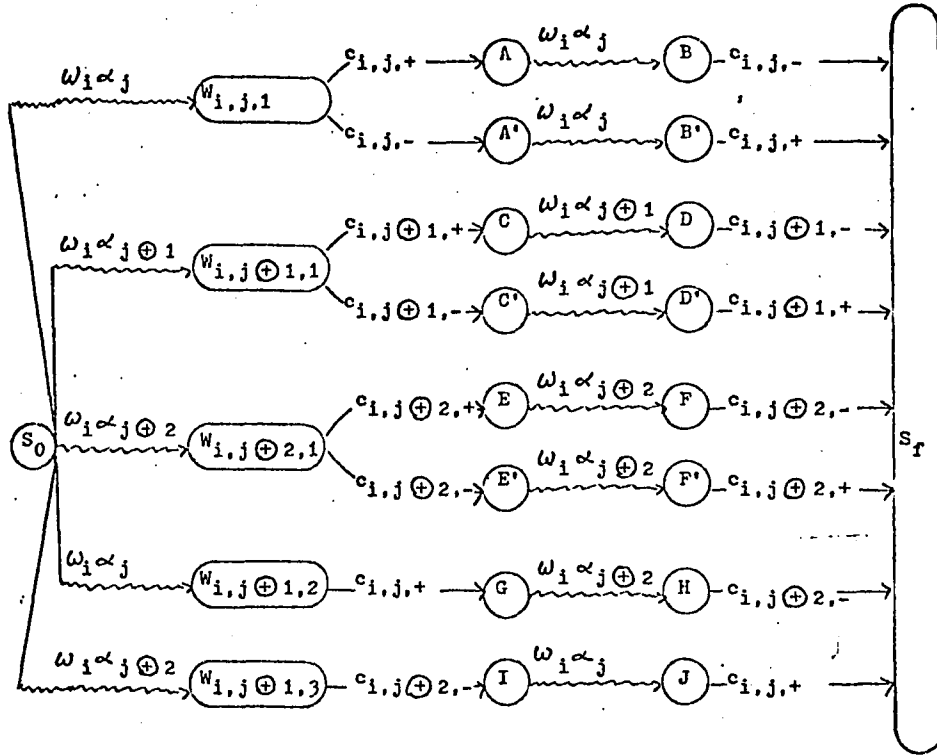
Treating sequences which do not parse as a series of Ds as ungrammatical (see discussion, p. 62) allows this  $DD^*$  interpretation of the data to be handled by a relatively straightforward nondeterministic modification of the finite state device of sec. 3. Suppose, as seems likely, that the appropriate well-formedness constraint to be imposed on constituent Ds is the semantic one of synonymy. This means that for each distance and direction indicating dance, the transition out of  $S_0$  must be to a state coding the semantic interpretation of that dance. In particular, it must not be the case (as it was in sec. 3) that all exits from  $S_0$  generating  $\omega_i \alpha_j$  are to the same state  $W_{i,j}$ ; a dance whose first constituent is  $\omega_i \alpha_j$  is probably not in general<sup>4</sup> thereby interpretable since it appears that it may

be followed by, among other things, the sequences  $c_{i,j,+}$   $\omega_i \alpha_j$   $c_{i,j,-}$  (giving  $\omega_i \alpha_j$  as the semantic interpretation) or  $c_{i,j,+}$   $\omega_i \alpha_j \oplus 2$   $c_{i,j} \oplus 2,-$  (giving  $\omega_i \alpha_j \oplus 1$  as the semantic interpretation). If the finite state device of sec. 3 is recast so that the path from  $S_0$  to  $W_{i,j}$  is part of only those dances semantically interpreted as  $\omega_i \alpha_j$  (in other words, so that only synonymous dances share a W-state), then the superdance claim may be accommodated by augmenting the device as follows: for each path linking a state  $S$  to  $S_f$ , introduce an identical path linking  $S$  to a new state  $SS$  which in turn exits to all states  $W_{i,j,k}$  via a path identical to that joining  $S_0$  to  $W_{i,j,k}$ , provided only that for some particular  $k$ ,  $W_{i,j,k}$  is indirectly connected to  $S$ . Thus if  $S$  is a penultimate state for generating a dance coding a location at  $\omega_4 \alpha_{16}$ , then  $S$  may exit to initiate any other dance coding that same location, i.e., by way of state  $W_{4,16,1}$  or  $W_{4,16,2}$ , etc.

The synonymy-corrected finite state device might look in part like (37), which would then be augmented as in (38), to meet the SD claim.

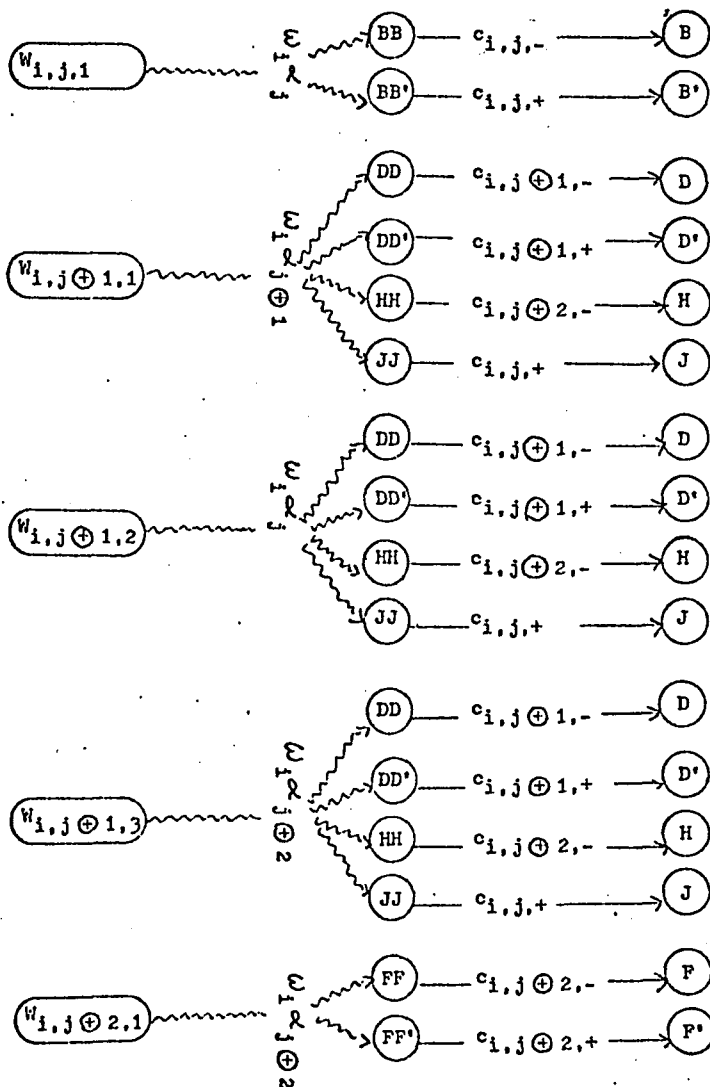
A decision to treat difficult-to-parse circuit orders (e.g.,  $\overline{LRRL}$ ) or within-SD dance-type mixtures as ungrammatical is also easy to accommodate within this theory. A finite state device can be constructed so that every distinct dance has a derivation path utterly distinct (except for the termination at  $S_f$  which all have in

(37)



(38)

Paths to be added to (37):



common), even as to initial state, from every other dance. Each path connecting a state  $S$  to  $S_f$  mandates an additional identical path from  $S$  to the initial state connected to  $S$ , and one from  $S$  to any other initial states generating a dance which may follow the first (e.g., might still want to permit the concatenation of  $S$ -dances differing only in the size or orientation of the curved runs).

The SD claim faces an additional hurdle, a more obvious one perhaps, in the shape of the marathon dance, during which (see pp. 68 - 70) "the dancer changes her directional information according to the movement of the sun . . . without ever observing the solar position directly through the entrance hole" (Lindauer 1967:95). It thus seems necessary not merely to permit, but somehow to ensure, the concatenation of messages not all synonymous. That is, to ensure that the angle (but not the wagging duration which presumably must be held constant at  $\omega_1$ ) is incremented at a rate and acceleration consonant with observable solar motion in the vicinity of that hive. This follows from the hypothesis that semantic interpretation consists in the lexical hookup of any  $\alpha$  and  $\omega$  values presented, followed by an averaging process giving the distance and direction to the goal;  $\alpha$  is read with respect to the upward or sunward direction, depending upon the dance's physical setting.

There is no obvious way of handling this phenomenon in a non-transformational fashion. Integral (over  $\alpha_j, 1 \leq j \leq n$ )

incrementing transformations, indexed to be used at particular seasons of the year and times of day for a specific location, are however guaranteed to exist--recall that transformations are of Type 0 and so, capture the notion of "effectively computable" processes: the fact that the bee learns to compute solar motion guarantees that an effective procedure (and therefore, a transformation) exists. The interesting question in this regard is that of deciding on a non ad hoc basis just where these transformations appear. That is, are they properly a part of a theory of dance language competence or of dance language performance?

If they are part of competence, a grammar of transformational complexity seems<sup>5</sup> indicated: a base component which (so far, at any rate) is finite state in design defines SD as all sequences of synonymous dances, which set is then filtered through a transformational component accounting for the incrementing. Each particular SD is realized as a large number of variably incremented sequences, as many as there are transformations in the dance language's syntax. It is pragmatics (performance theory at the semantic level) which is responsible for characterizing the spatial and temporal conditions making a particular incrementing of the SD appropriate; it is pragmatics which is responsible for deciding on the basis of the context of a given utterance of some SD, which of the finite range of transformations must be utilized. Just as pragmatics

determines on the basis of context which of "Happy Birthday!", "You're welcome.", ... (all sentences freely generated by the grammar) is appropriate to a given occasion. The grammar sketched in this paragraph is formally quite similar to that of English; the primary structural difference lies in the base component's being of Type 3 for the bee and Type 2 for man.

If, however, they are part of performance, relegated to pragmatics, the structural similarity between the two species-specific grammatical theories vanishes. The grammar (again, so far) is finite state, SD defined as any sequence of synonymous dances. The spatial and temporal context of an utterance of a given SD is accessed by pragmatics which selects from its repertoire the appropriate incrementer, whose probable transformational character, now that it is outside the grammar, is no longer of much significance.

So, once more, are these transformations syntactic or pragmatic in nature? To put this question another way, is the learning which leads to incrementing linguistic learning? Is there a language acquisition process in the honey bee? Or is the honey bee's learning restricted to learning about the world, rather than learning about its language?

The evidence, such as it is, clearly favors the view that this learning is not linguistic, that the dance

language is not in any way acquired. In other words, that solar incrementing in the marathon dance is properly treated as pragmatic.<sup>6</sup>

Three observations suggest this point of view. The implication is that the investigation of learnable incrementings is not linguistic, but more generally cognitive in nature. This is not to say that such studies are uninteresting or valueless; it would be particularly interesting to determine the characteristics of the innate component which provides for the extrapolation from exposure to a portion of an incrementing curve to an approximation of the full curve,<sup>7</sup> whether there are developmental pressures on the learning of solar motion, the response to hive displacements severe enough to change the perceived pattern of solar motion,<sup>8</sup> what constraints there are on learnable incrementings. The evidence against attributing this learning to the linguistic system establishes three properties of the marathon dance incrementing process which suggest that the wider cognitive interpretation is to be preferred.

First, incrementing is not restricted to dance production behavior, but is also reported to be a response during the training to a feeding site, which precedes all the dance language experiments. Initially, the feeding station must be inched away from the hive entrance in order not to lose foragers, but the distance between halts may be gradually increased, with von Frisch 1967a:17 claiming

further that in time, the foraging group was actually able to anticipate his relocation of the feeding table. Assuming that there is in fact a gradual statistical bias favoring what may be described as forward error at the expense of leftward, rightward, etc. error in feeding table search behavior, this is impressive. Such a result would only be explicable as due to olfactory cueing if the behavior were confined to downwind movement of the feeding station. The success of the training method described in Gary and Witherell 1971, which advises (on p. 449) that "[f]eeders should be placed at multiple stations, spaced apart at increasingly longer intervals, and oriented in 1 direction from the hive," is additional evidence of incrementing ability in this non-linguistic context. It is noteworthy that both of these are distance incrementings. Attempts at provoking a non-linguistic directional incrementing response in the bee have been confined to the use of a portable feeder shifted in mockery of the sun's diurnal path so that the difference in azimuth between the goal and the sun was kept constant despite the sun's movement; von Frisch 1967a:207-10 reports that these "motionless sun" experiments confused all but Lindauer's cellar bees (see n.7).

New and New 1962 gives evidence of variable incrementing ability within a hive presented with the sun at the perceptual zenith. While true disoriented dancing did occur, many dancers indicated (on the diffusely lit vertical

dance floor where gravity provides a frame of directional reference despite the near-zenith sun) the goal's position relative to a solar calculation: a two-way extrapolation which "between successive memorized positions turns the dances through an angle proportional to time" (New and New 1962:287). Genetic factors are hypothesized to account, at least in part, for the clockwise vs. counter-clockwise turning pattern bias of individual dancers. What is more interesting, however, in the current context, is the fact that there is individual variation with regard to orientation maintenance during a site advertisement, with some few dancers producing disoriented dances. Thus, there is within-hive variation in the quality of the extrapolation itself, variation as to whether or not the direct sensory monitoring of solar azimuth is supplemented by this two-way extrapolation filling in the blanks, so to speak. If the acquired solar incrementing functions are to be considered part of competence it must be supposed that there is competence variation within the hive.<sup>9</sup>

Finally, since the incrementing transformations recapitulate the extrapolation of perceived solar motion, they necessarily vary with the location of the hive. If the transformations are a syntactic part of dance competence, then each (sufficiently different) possible hive location defines a new dialect of the dance language. In other words, the innate properties of the dance language are (dance)

linguistic universals; the transformations are (dance) language particular. Carniolan bees in different locations do not literally share a language. What's more, it follows that it is purely coincidental that the transformations discriminating the different Carniolan languages are always selected from a set of functions describing solar motion as perceived from various terran locations.

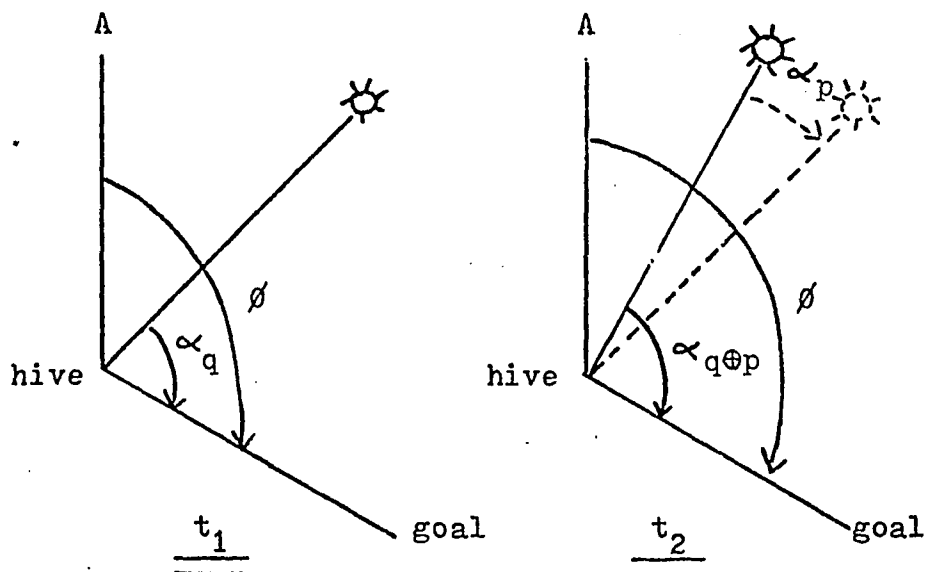
Now, while admittedly none of this is conclusive it certainly acts to undermine this attempt to motivate transformational complexity for the dance language. The best supported syntax for this DD\* version of the SD claim, again, is a finite state sequencing of synonymous dances which is subject to a performance incrementing sensitive to spatial and temporal dance site contextual parameters.


The precise machinations of the pragmatic incrementing transformations are then in question. As a part of performance theory, the process may best be examined in the light of the probable perceptual substrate for the learning of solar incrementing. Since solar azimuthal change can only be recognized in relation to some fixed feature A of the hive's environment, it is worth considering the hypothesis that solar tracking is done with respect to that feature. The fact that A is a fixed<sup>10</sup> environmental feature and yet it is directional indication with respect to a moving sun that evolution selected for, suggests that A is in general not a possible orienting force in flight. That

is, there is an immediate teleological argument against A's being, for example, the earth's magnetic field<sup>11</sup> which, since constant and (apparently, at least) equally perceptible on the hive-goal flight as at the hive itself, would be a directional reference axis superior to the diurnally and seasonally variable solar azimuth. The obvious alternative is that A is a landmark in the vicinity of the hive, in which case it may well differ from bee to bee. If for each dancer there is a landmark A to play such a role, then the synonymy condition may be drawn with respect to A, as follows.

Consider the time interval  $(t_1, t_2)$ , during which the sun moves, for example, counterclockwise a quantity  $\alpha_q$ , changing the danced direction from  $\alpha_p$  at  $t_1$  to  $\alpha_{p \oplus q}$  at  $t_2$ , as in (39). By the dancer's ability to do solar

(39)





incrementing is meant her ability to sensorily monitor, and extrapolate from previous sensory monitorings of, the difference in azimuth between A and the sun. The perceived/calculated A -  is subtracted from the constant angle  $\emptyset$  between A and the goal, producing the variable danced directional value, as in (40). Now if  $\alpha$ -interpretation,

$$(40) \quad \frac{\text{time} \quad (A\text{-to-goal}) \ominus (A\text{-to-sun}) = \text{sun-to-goal}}{\text{danced value}}$$

$$t_1 \quad \emptyset \quad \ominus \quad (\emptyset \ominus \alpha_q) = \alpha_q$$

$$t_2 \quad \emptyset \quad \ominus \quad (\emptyset \ominus \alpha_{q \oplus p}) = \alpha_{q \oplus p}$$

where  $x \ominus y = z$  if and only if  $x = y \oplus z$ .  
(See sec. 3's n.5 for a definition of  $\oplus$ .)

unlike the lexical lookup sketch assumed thus far, consists in transposing the angle with respect to the sun's azimuth  to one with respect to her A, whatever it is, then the synonymy condition stands. The synonymy condition stands, not because solar incrementing transformations are to be found in pragmatics (performance theory) rather than syntax (competence theory) but because there is a performance translation linking direction with respect to  to the competence coding of direction with respect to A; solar incrementing is an integral part of that translation mechanism, but now has no linguistic status for either the competence or the performance analysis. Again, a finite state syntax as sketched will serve as the model.

This A-analysis derives support from experimental evidence that in some circumstances natural landmarks are


preferred to the sun as an in-flight orienting force. The paradigm case is a hive to goal flight line parallel to a row of trees; experimental displacement of the hive from an environment with a north-south line of trees to one with an east-west line of trees led the foragers to seek the goal in a westerly as opposed to a southerly direction, ignoring the solar compass (result due to von Frisch and Lindauer 1954). Subsequent experimentation (discussed in Lindauer 1967:126-30 and von Frisch 1967a:339-47) designed to isolate the landmark properties that contribute to preferential treatment in orientation argue that it is an unbroken linear display (of trees, shore, road, etc.) which is the determinative feature; single trees or groups of trees did not lead the bees astray. It is a curious property of this work that is is visual similarity before and after hive displacement that is controlled for; the presumption is that the landscape is visually encoded in the bee. This seems an unwarranted assumption given the great importance of olfactory information for this species, and calls into question the "unbroken linear display" conclusion of that work.<sup>12</sup>

At the same time, of course, it calls into question the assumption that A is a localized environmental feature or landmark. This notion is not logically dependent upon a visual landscape coding but it is certainly encouraged by such a conception. The observation that the sun's azimuth,

☀, shifts diurnally can, in that case, only be made with respect to other stationary landscape feature(s) visually coded. The fact that this is viewed as a visual competition between ☀ and something else encourages a "frictionless plane" idealization of that something else as a point, A. If, however, the competition is between a stationary, dynamic olfactory coding and an overwhelming, obtrusive visual stimulus ☀ diurnally varying in orientation with respect to that olfactory map, this idealization to a point, A, is suddenly suspect. There is no reason to assign olfactory primacy to any feature and indeed, some reason against it, given that the odor-intensity at the different sites in the vicinity of the hive varies with the time of day according to the species of plant growing at that location. (Different plant species vary as to the time of the day they open their flowers.)

No experimental test of the olfactory/visual landscape coding hypothesis is possible, forcing a direct investigation of the A-analysis, parodying the "motionless sun" experiments (see p. 142). That is, various landmarks, olfactorily or visually striking, must be moved (one in each experiment in the series) in synchrony with the sun for a period of days, and the solar incrementing ability of those bees who should have been learning solar motion during that period subsequently investigated. A substantial number of apparent "cellar bee" dance performances

confirms the A-analysis. At this present, the A-analysis seems implausible in view of the olfactory landscape coding possibility. As will be shown, however, it has considerable descriptive force and is therefore pursued here as one of three accounts of solar incrementing.

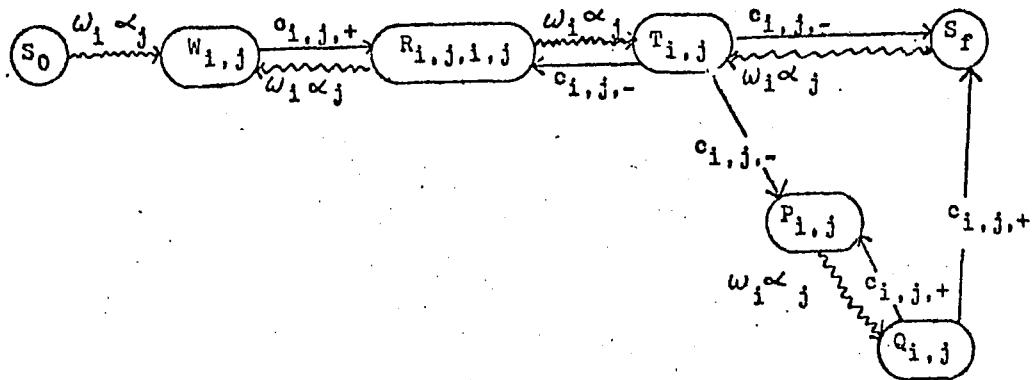
The argument in favor of the A-analysis is not unlike that proposed in sec. 2.11 (see (8)) for the pragmatic selection of an axis in the case where light and gravity are contextually induced competing modes of directional orientation. At that time, the discovery that the interpretation of the danced angle in terms of the sun or in terms of the upward direction was prohibitively complex and, less importantly, that there was no basis for a decision between them, led to the hypothesis of a new directional reference point, the contextually weighted average of the solar and upward directions. In the present case, the A-analysis provides a similarly simpler alternative with respect to directional averaging during the marathon dance. If the linguistic encoding of direction is a translation between direction with respect to A and direction with respect to , then it is not necessary to decide what solar position is taken as dominant in a marathon dance series: all solar positions are of equal status, all directional indications are interpreted with respect to A. A non-landmark analysis, on the other hand, is drawn into the uncomfortable position of having to favor in (39), either

the  $t_1$  (averaging  $\alpha_q$  with a  $t_2$ -perceived  $\alpha_{q \oplus p}$  incremented-corrected to  $\alpha_q$ ) or the  $t_2$  (averaging  $\alpha_{q \oplus p}$  and a  $t_1$ -perceived  $\alpha_q$  incremented-corrected to  $\alpha_{q \oplus p}$ ) perception, incrementing the averaged value ( $\alpha_q$  if  $t_1$  is favored,  $\alpha_{q \oplus p}$  if  $t_2$  is favored) to the new value appropriate at takeoff time  $t_3$ . The A-analysis is faced with no such unmotivated choice.

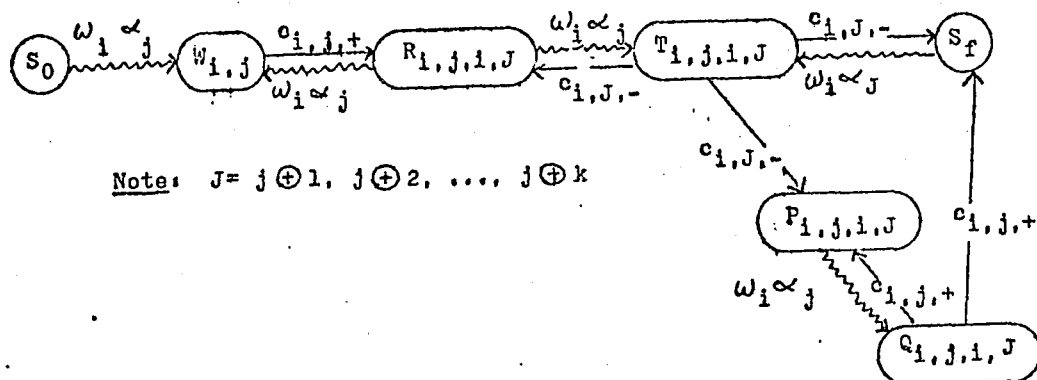
This in broad outline is the finite state response to the SD claim interpreted as a message series, each message unit D having the syntactic structure specified in sec. 3; in other words, this is the observational adequacy response to the SD  $\rightarrow$  DD\* version of the SD claim. The "no-D" version of the SD claim is less explicit and so, more difficult to respond to. In effect this no-D analysis extends the admitted "sloppiness" of the round dance to the other dance forms, treating the non-parsable circuit orders (e.g., LRRRL) as grammatical as well as (possibly) acceptable. What is not explicit, of course, is the degree of sloppiness to be tolerated. Presumably, LLL is still ungrammatical as a full SD, but is LLLLLLR grammatical?

Assuming for the purposes of illustration that the criterion of grammaticality is at least one L and at least one R circuit (in either order), the tailwagging portion (12) of sec. 3's finite state account should be revised as in (41). The tailwagging transitional portion (14) of sec. 3's finite state account is similarly revised, as outlined in (42).

(41)



(42)



Revisions along these lines are essentially conservative, carefully allowing only one dance type (and indeed, in the above case, only one specific degree of angular spread in the transitional dance) per SD. However, if semantic interpretation is straightforward averaging across all indications to a specific  $\alpha$  and  $\omega$  value, then the concatenation in (42) must be further constrained to mention values which together average to some  $\alpha_j$  (and not  $\alpha_j \oplus 1/2$ )

for example), these being cognitive (competence) categories).

The situation is much simplified, of course, if sequential repetitions of a particular circuit do not weight its contribution to the average. However, in that case, it is difficult to see what evidence there might be for the no-D analysis in the first place. That is, in the event that the two SD's (43) and (44) both indicate a goal

$$(43) \quad \omega_i \alpha_j \quad c_{i,j,+} \quad \omega_i \alpha_{j \oplus 12} \quad c_{i,j \oplus 12,-}$$

$$(44) \quad (\omega_i \alpha_j \quad c_{i,j,+})^2 \quad \omega_i \alpha_{j \oplus 12} \quad c_{i,j \oplus 12,-}$$

$$(\omega_i \alpha_j \quad c_{i,j,+})^4 \quad \omega_i \alpha_{j \oplus 12} \quad c_{i,j \oplus 12,-}$$

situated in the direction  $\alpha_{j \oplus 6}$ , the best analysis of the six-L, two-R superdance of (43) would seem to be in terms of "stuttering," i.e. a performance problem. There would be no need to treat it as grammatical, but only (perhaps) acceptable. It seems that giving the no-D analysis the benefit of the doubt involves allowing all danced maneuvers with semantic content to contribute to the SD's semantic interpretation. The problem of ensuring that the average so produced is of the proper form would be even further complicated by allowing mixed-dance concatenations (such as, for example, (45)), which is more in the spirit of the no-D analysis than the very conservative view of (41) and (42). Presumably, at least some  $\omega$ -variation would be

$$(45) \quad \omega_i \propto_j \quad c_{i,j,+} \quad \omega_i \propto_{j \oplus 4} \quad c_{i,j \oplus 4,-} \quad \omega_i \propto_{j \oplus 4} \\ c_{i,j \oplus 4,-} \quad \omega_i \propto_{j \oplus 4} \quad c_{i,j \oplus 4,+}$$

permitted across an SD, provided again that the average is to  $\omega_i$  for some  $i$  and not, for example, to  $\omega_i \oplus 1/4$ .

Now, the combined effect of these criteria (at-least-one-L, at-least-one-R, indefinitely long SD's, and the integral constraint on averaging) for grammaticality is to preclude a finite state treatment; what's more, it is not at all obvious just how these sequences might be specified syntactically, even given transformational capacity. This being the case, it is of course a temptation to shunt the averaging condition into a semantic component of the dance grammar, freely generating via a finite state syntax all at-least-one-L, at-least-one-R sequences which show the requisite degree of  $\omega$ -agreement (probably constant across dance types, with all wagging runs in a SD restricted to a range of, for example, three adjacent  $\omega$ -values) and  $\propto$ -variation (this dependent on dance type or, in other words, wagging duration--small  $\omega$ -values correlate with a large  $\propto$ -range, etc.). Those grammatical SD's which fail on averaging to give integral  $\propto$  and  $\omega$  values are subsequently marked semantically anomalous.

The no-D analysis, for which it must be emphasized no positive evidence exists nor is likely to be discovered, is discussed in such detail in order to clarify just how

severe a problem is posed by a cognitively based digital analysis of the dance language. To a great extent, this problem was masked by the nature of the idealizations which were perhaps somewhat more natural than those being considered here; it must be emphasized just how great a gap was thereby made between the dancer's competence and performance, again necessarily without experimental evidence.

Before turning to a discussion of the descriptive adequacy of the finite state approach, it may be helpful to consider the nature of the arguments that have been used toward a demonstration that transformational capacity is required for an observationally adequate theory of human language,<sup>13</sup> as a preliminary to an outline of what would constitute analogous dance language data. This is intended to justify the departure at this time to the consideration of the descriptive adequacy of the finite state approach.

As was indicated at the beginning of this section, the question of the observational adequacy of a finite state grammar only arises for a language consisting of an infinite set of sentences. The argument<sup>14</sup> that English is infinite draws heavily upon the competence/performance distinction, relegating the explanation for the general increasing-with-length difficulty with such sentence sets as (46) and (47) to a psychological (i.e., performance) rather than a linguistic (i.e., competence) theory.



more easily comprehended than (49).<sup>15</sup> Other arguments include the comprehension . asset derived from the use of pencil and paper in such cases, this clearly being an aid to the subject's memory (performance) and not to his internalized linguistic system (competence).

From the grammarian's point of view, the chief advantage of the indicated idealization of English to a suitable infinite set is the introduction of recursion into the grammar--the accounting for such syntactic patterning as that illustrated here without imposing rule repetition (syntactic complexity) limits. A finite state account of a finite set, whether for English or for the dance language is no more revealing structurally than a list of the grammatical strings would be. It is essentially the proper competence infinite-izing of the dance language which is at issue here and in sec. 4.2.

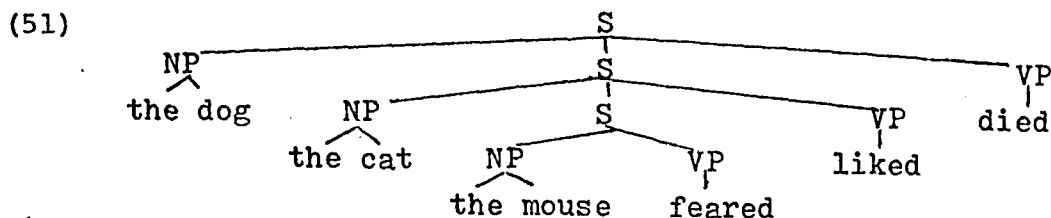
The best version (due to Corstius, as reported in Levelt 1974b:25-26) of the argument that English, and therefore human language, is not regular (i.e., observationally amenable to a finite state treatment) is of the reductio ad absurdum form, and is based on a variant of the English subset given in (46). Assume English is regular and consider its intersection with the regular set  $S = \{(\text{the oyster})^n (\text{split})^m / n, m \geq 1\}$ . Since the intersection of a regular set with another set has the formal properties of the other set, the intersection of English and  $S$ , by

hypothesis  $\{(the\ oyster)^n\ (split)^n / n \geq 1\}$ , must if English is a regular set, be a regular set. However, it has been shown that such sets are not regular. Therefore, if the data is as outlined, it follows that English is not a regular set.

The "proof" is usually followed by a demonstration of the descriptive advantages of a context free account on the order of that given in (50), which generates tree

- (50)  $S \rightarrow NP\ S\ VP$        $NP \rightarrow the\ mouse,\ the\ oyster,\ the\ dog,\ the\ cat,\ \dots$   
 $S \rightarrow NP\ VP$        $VP \rightarrow liked,\ feared,\ split,\ died,\ \dots$

diagrams of the form (51), tree diagrams being in effect

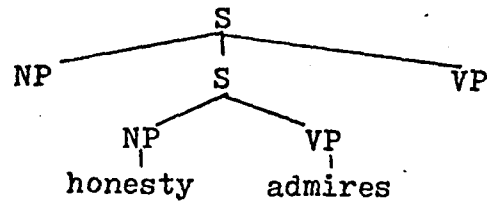


derivational histories of a context-free generated terminal string. (51) sets the stage for the discovery of a relationship between the NP ("noun phrase") and VP ("verb phrase") entered by the same rule (e.g., the cat and liked, the mouse and feared). And, indeed, the NP functions as the subject of the VP with which it is paired. That is, only those NP's which may serve as the subject of a given VP in a simple sentence may be entered with it by an S rule. So, since (52) is ungrammatical, so is any string associated

with a completion of (53). The fact that the verbs embedded

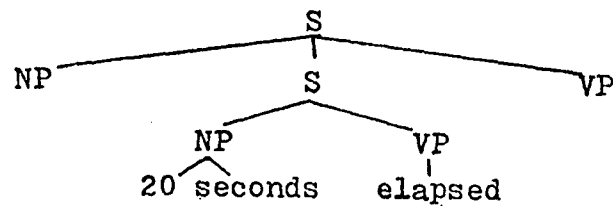
(52) "Honesty admires..."

(53)



by this rule must be transitive ones in simple sentences (since elapse is intransitive, only ungrammatical strings are associated with a completion of (54)) and, moreover,

(54)

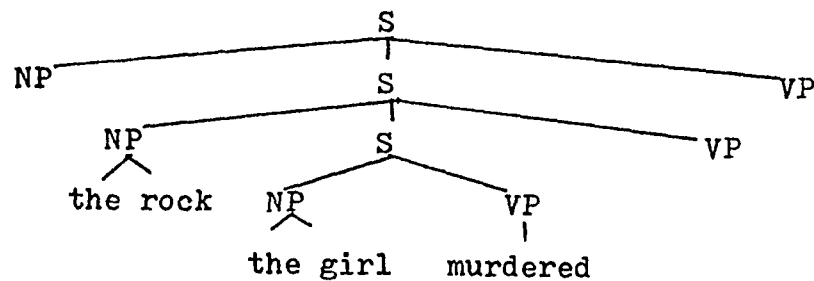


must be able to have as object the particular subject of the next highest level (so since (55) is ungrammatical,<sup>16</sup>

(55) \*The girl murdered the rock.

so is any string associated with a completion of (56))

(56)



remains unexplained, demonstrating that this is a descriptively inadequate analysis.<sup>17</sup>

The SD data is sufficiently vague to permit a center-embedded version, i.e., one with the property just pointed out for English. Suppose that the operative condition defining a grammatical SD is neither that it parse into a sequence of entities (D's) with a specifiable syntactic structure nor that it have at least one L and R, but rather that it be an arbitrarily long series of left circuits followed by a precisely equally long series of right circuits, as in (57).<sup>18</sup> Then the argument that English is not

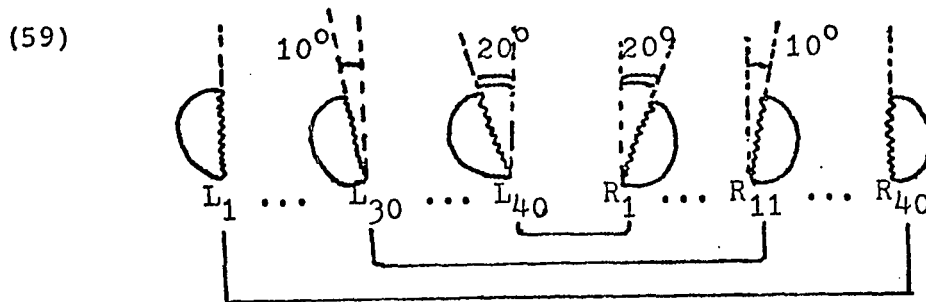
$$(57) \quad \left\{ ( ( \text{☺} )^n ( \text{☹} )^n / n \geq 1 \right\}$$

regular may be mimicked in a straight-forward way, showing that the dance language is not regular by intersecting it with the set  $\left\{ ( ( \text{☺} )^n ( \text{☹} )^m / n \geq 1, m \geq 1 \right\}$ . A descriptive advantage to a context free analysis on the order of (58) might be adduced if the pairings thereby introduced

$$(58) \quad \begin{array}{l} D \rightarrow ( \text{☺} ) D ( \text{☹} ) \\ D \rightarrow ( \text{☺} ) ( \text{☹} ) \end{array}$$

were shown to be linguistically significant, as by means of the output of the solar incrementing pragmatic process. That is, the given context-free analysis would have a definite descriptive advantage if the series  $( ( \text{☺} ) ( \text{☺} ) ( \text{☺} ) ( \text{☺} ) \dots )^n ( ( \text{☹} ) ( \text{☹} ) ( \text{☹} ) ( \text{☹} ) \dots )^n$  were realized in performance as a series of sun-corrected left circuits followed by a same-size series of right circuits which recapitulates the effect of the sun's motion in reverse. Thus, supposing

that the sun's accelerating counterclockwise motion during an 80-circuit marathon tailwagging dance forces a  $10^\circ$  clockwise shift from the first left circuit's ( $L_1$ 's) danced direction at  $L_{30}$  and another at  $L_{40}$ , it is required that the right circuit corresponding to  $L_{30}$  (viz.,  $R_{11}$ ) show a  $10^\circ$  clockwise perturbation and that corresponding to  $L_{40}$  (viz.,  $R_1$ ) a  $20^\circ$  clockwise perturbation from the final right circuit  $R_{40}$  identical to  $L_1$ , as in (59). While this



interpretation of the SD claim is admittedly unwarranted, this "descriptive advantage" certainly misconstrues the data. At best an observational adequacy argument might be mounted.

Equally unsupported versions of the SD claim might be constructed using any of the other patterns known not to be amenable to a finite state treatment. The fact that human language has not been shown to display anything comparable is irrelevant; the dance language would in any case be shown not to be regular.

The best version (again due to Corstius, as reported in Levelt 1974b:31-32) of the argument that English,

and thus human language, is not context free, is again of the reductio ad absurdum form and is based on the data of (47). A (by assumption) context-free English is intersected with the regular set in which two occurrences of respectively link coordinate structures of various sizes; the result must, if English is context-free, be context-free. However,  $\{x^n \text{ gave } y^n \text{ respectively } z^n \text{ respectively} / n \geq 1\}$  is known not to be context-free. Thus, if the data is as outlined in (47), English is not context-free.


None of the other patterns known not to be amenable to a context-free treatment has been shown to naturally extend to English.<sup>19</sup> None at all appears to extend to the dance language of the honey bee.


Despite the label, it is particularly hard to envision an SD version having the mirror-image property, i.e. the property that all strings, of which there are infinitely many, divide into two strings of equal length, each the reverse of the other. This is not the sense of "mirror image" in sec. 3 n.3. In particular, the tailwagging dances of (60) and (61) do not reveal a mirror pattern of the



$$(60) \quad \left( \begin{array}{c} \uparrow \\ \downarrow \end{array} \right) \left( \begin{array}{c} \uparrow \\ \downarrow \end{array} \right) \quad \omega_i \propto_j c_{i,j,+} \quad \omega_i \propto_j c_{i,j,-}$$

$$(61) \quad \left( \begin{array}{c} \uparrow \\ \downarrow \end{array} \right) \left( \begin{array}{c} \downarrow \\ \uparrow \end{array} \right) \quad \omega_i \propto_j c_{i,j,-} \quad \omega_i \propto_j c_{i,j,+}$$

indicated kind; (60) neither displays an internal mirroring (the mirroring of  $\omega_i \alpha_j c_{i,j,+}$  in (60) would be  $c_{i,j,+}$

$\omega_i \alpha_j$  , giving the clearly ungrammatical \*  ) nor does (61) mirror (60). (The mirroring of (60) would be a subsequent  $c_{i,j,-} \omega_i \alpha_j c_{i,j,+} \omega_i \alpha_j$  ,

giving the, again, clearly ungrammatical \*  ,

danced in the order   ). Given a dance notation in terms of wagging and curved runs, mirroring requires a single, centered waggle-waggle or curve-curve sequence in each of the infinitely many SD's.

Abandoning this notation for a circuit notation  $C_{i,j,+}$  (for what has previously been described as  $\omega_i \alpha_j c_{i,j,+}$  ) incidentally forcing a horrendous holistic  $S_{i,j,p,q,+,r,q}$  (for  $\omega_i \alpha_j c_{p,q,+} c_{r,q,-}$ ),  $S_{i,j,p,q,-,r,q}$  (for  $\omega_i \alpha_j c_{p,q,-} c_{r,q,+}$ ) notation on the S-dance, gives an interesting, though again quite suspect mirroring version of the SD claim. The mirroring version would then define the tailwagging SD as in (62), namely as a series of LR alternations followed by a series

$$(62) \quad \left\{ (C_{i,j,+} C_{i,j,-})^n (C_{i,j,-} C_{i,j,+})^n / n \geq 1 \right\}$$

of RL alternations, in other words a LR alternating series

beginning and ending with L and with a double R dead center. On the whole, this is a less objectionable interpretation than that which allows a freer circuit order before the execution of the double R and requires its exact reversal afterward. Neither of these, however, can be said to be a particularly attractive competence idealization. Nor is it a more attractive approach to an idealization of some other dance type.

In sum, the finite state account is least threatened (on observational grounds) by the dancer's solar incrementing capacity, which it seems reasonable to eliminate from the grammar either by way of the A-analysis or by means of a pragmatic treatment. A no-D analysis taking SD as the syntactic domain suggests a semantic integral-averaging constraint on the finite state syntactic output. There is no obvious extension from the argument pattern used to demonstrate the failure of finite state linguistic theory for human language to the dance language.

The descriptive adequacy of the finite state account is now in question.

#### 4.1.2 Descriptive Adequacy

Historically, it was the goal of descriptive adequacy which motivated the theory of transformational grammar. Later, arguments were mounted purporting to demonstrate that neither finite state nor context-free phrase structure

grammar were observationally adequate, but initially the debate centered on the formal specification of a linguistic theory able to capture the linguistically significant generalizations of human language. Perhaps the most obvious approach to the contrastive analysis of the dance language and human language is provided by a discussion of the use which might be made of typical human language operations within the dance language grammar.

Transformational grammar uses context-free phrase structure rules to hierarchically relate category symbol concatenations. The lexicon consists of a list of the language's morphemes, each labelled as to the particular syntactic constellation in which it can appear; this labelling ranges from stipulating the category of the item (so, old is labelled an Adjective) to more global properties of the environment (so, the Verb know must be prevented from occurring with the Progressive -ing suffix, as in \*He is knowing the answer.). The addition of lexical items to the hierarchically arranged category symbol concatenation generates what is known as the sentence's deep structure, which is then siphoned through a series of transformations which delete, insert, or move lexical material dominated by particular categorial configurations. The output here is known as the sentence's surface structure. The standard theory holds that transformational rules are semantically inert, not able to change

the meaning; it is the deep structure which is the input to the standard theory's rules of semantic interpretation, while the phonological rules, which concern patterning at the phonetic level, apply to the surface structure.

Fodor (1966:292-93) presents several behavioral criteria for the interjection of hierarchical structure into the discussion of the nonverbal behavior of infra-humans. The discovery that certain substrings of a given string of behavior are more easily shaped or extinguished, more frequently flanked than interrupted by pauses, or more prone to displacement than others is, he argues, to be captured by associating them derivationally, as is a substring of innate behavior which is affected as a unit by cross-breeding. It is an argument for establishing hierarchical structure by establishing multiple analytical levels as functioning within the performance system, "demonstrated by the fact that operations of extinction-relearning, generalization, and the rest of the paraphernalia of behavior change can operate not only on the elements but also on the units into which the elements are bracketed" (Fodor 1966:292). Two analytical levels must be established before a dominance relation can be presumed to hold between them.

Evidence for an  $\omega$ -feature, therefore, is evidence that the elements of  $\{\omega_i \alpha_j / 1 \leq j \leq n\} = \{\omega_i \alpha_1, \omega_i \alpha_2, \dots, \omega_i \alpha_n\}$  for any given  $i$ ,  $i \leq i \leq m$ , have a common performance

property à la Fodor 1966. It must be shown that all wag-  
gling runs which agree along the hypothesized feature are  
set apart by some behavioral factor. In this case, there  
are two.

Assuming the distance semantics which is traditional  
in the dance language literature, the fact that all dances  
whose wagging runs are of a particular duration (say,  $\omega_i$ )  
are identically reclassified with respect to distance  
representation under interracial crossbreeding is of crucial  
significance--the only possible explanation for an identical  
crossbreeding output of this sort is an identical cross-  
breeding input, in other words an  $\omega$ -feature. The argu-  
ment is impeccable, the only weak link the assumed distance  
semantics; a semantics in terms of the percentage of the  
flight range (see sec. 2.4.2) strips interracial cross-  
breeding of any semantic effect and thus this argument of  
any force. It is therefore necessary to search for yet  
another, less equivocal, performance reflection of this  
parameter.

What is needed, again, is evidence that dances with  
wagging runs drawn from the set  $\{\omega_i \propto_j / 1 \leq j \leq n\}$  for  
a given  $i$ ,  $1 \leq i \leq m$ , are in some respect identically  
accessed by the production or comprehension mechanism. This  
is provided by the investigation of the angular scatter pro-  
duced during an SD, which Gould 1975c reports in passing  
uniformly varies inversely in magnitude with the wagging

duration value  $\omega_i$ .<sup>20</sup> That is, the observed correlation between the duration of wagging/sound production and the deviation in danced angle  $\alpha$  between successive circuits for the Y-shaped transitional dances is not peculiar to that dance type, but recurs across message repetitions in the other wagging cases. The psychological reality of the  $\omega$ -feature may therefore be adduced not only for those values of  $i$  peculiar to the Y-shaped transitional dances but in fact for all  $i$ ,  $1 \leq i \leq m$ .

Evidence for an  $\alpha$ -feature is, in like fashion, evidence that the elements of  $\{\omega_i \alpha_j / 1 \leq i \leq m\} = \{\omega_1 \alpha_j, \omega_2 \alpha_j; \dots, \omega_m \alpha_j\}$  for any given  $j$ ,  $1 \leq j \leq n$ , have a common performance property. It must be shown that all wagging runs which agree along the hypothesized feature  $\alpha$  are identically affected by some performance factor. Now, as previously argued (pp. 140-50), the incrementing-with-perceived-solar-azimuthal-change may well be a performance operation. It is unfortunate therefore in this regard that the particular incrementing imposed by a spatio-temporal context  $c$  is not a function of the particular  $\alpha$ -value in question. However, whatever the azimuthal difference between the goal and the sun (from  $\alpha_1$  to  $\alpha_n$ ), a three-unit  $\alpha$  shift in solar azimuth will shift that value three units (clockwise from  $\alpha_4$  through  $\alpha_n$  and on to  $\alpha_3$ , counterclockwise from  $\alpha_{n-2}$  through  $\alpha_n$  to  $\alpha_{n-3}$ ). Since the incrementing is not  $\alpha$ -value particular,

it can provide no support for an  $\alpha$ -feature.

What is  $\alpha$ -value particular is residual misdirection, the systematic deflection toward the horizontal and vertical axes which afflicts dancing on a vertical surface, the direction and extent of the affliction depending upon the relationship of the given  $\alpha$ -value to these primary axes (see (7)); since residual misdirection is peculiar to dances executed on a vertical dance floor, it is of the requisite physical, performance quality.

The  $\omega_i \alpha_j$  feature representation of the wagging run is thus behaviorally supported.<sup>21</sup> The hierarchical structure of the full terminal string in relation to the demonstrated  $\omega_i \alpha_j$  substring can now be investigated.<sup>22</sup>

Evidence for a category W, the wagging run, is evidence that the elements of  $\{\omega_i \alpha_j / 1 \leq i \leq m, 1 \leq j \leq n\}$  have a common performance property. The above-described (pp. 100-101) evidence that dances begin with wagging runs is suggestive, but no more. The observation (Esch 1964, as reported in von Frisch 1967a:105-106) that introducing a light source into a dark hive during a by-gravity dance produces an immediate by-light wagging run of the appropriate duration is more conclusive. The buzzing run (see p. 18) is perhaps also best viewed as a displaced wagging run rather than an additional dance type.

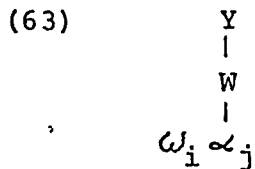
Evidence for a category C, the circuit, is evidence that the elements of  $\{\omega_i \alpha_j \quad c_{i,j,+} / 1 \leq i \leq m, 1 \leq j \leq n\}$  U

$\{\omega_i \alpha_j \ c_{i,j,-} / 1 \leq i \leq m, 1 \leq j \leq n\}^{23}$  have a common performance property. Evidence for a category L, left circuit, involves the first of the two unioned sets; evidence for a category R, right circuit, involves the second. For none of these supposed categories is a performance motivation obvious. In each case what is required is a demonstration that wagging runs followed by a return run (clockwise, counterclockwise, or either, for R, L, and C, respectively) are treated by the production or comprehension model differently from other wagging runs. Now since the wagging run by all accounts codes the information, this is unlikely to be the case. If solar incrementing were restricted to circuited dance types, it would argue for a category C, which would also be supported by a circuited dance-type restriction with respect to the new wagging run provoked by a light source interrupting dark-hive dancing. Nothing in the literature suggests either behavior is peculiar to circuited dance types. In an exactly analogous fashion, if solar incrementing were confined to left (right) circuits of circuited dance types or interrupting a dark-hive dance with a light source uniformly provoked by a left (right) circuit, this would be evidence for a category L (R).

Thus far, then, there is performance evidence of the requisite sort for the  $\omega_i \alpha_j$  representation of a substring of the behavioral strings in question, and for a category

W dominating  $\omega_i \alpha_j$ , no evidence for a higher category C, L, or R. The racially linked form of the transitional dance, even if only as a statistical preference, clearly points to a higher dance type category Y, identified as the Carniolan dance behavior which alternates genetically with some Italian sickle configuration. Y might be, following (14) and ignoring the SD claim,  $\{ \omega_i \alpha_j \ c_{i,j,+} \ \omega_i \alpha_j \oplus x \ c_{i,j} \oplus x,- \ / \ 1 < i_1 \leq i \leq i_2 < m, \ 1 \leq x \leq k = n/6 \}$ , where  $i_1$  and  $i_2$  define the range of wagging/sound production duration values appropriate to the dance type.

We are therefore led, according to Fodor 1966, to posit structural descriptions for the dance language of at least the complexity illustrated by the branch in (63).



Although Lockett and Altmann 1968 are regrettably vague on the subject of how to apply their language design feature criteria<sup>24</sup> to infrahuman behavior, evidence for a bracketing such as that in (63) appears to be in part what is demanded by their criterion "duality of patterning." What is additionally at issue, however, with regard to this design feature is the semantic characterization of the hypothesized terminal node(s); it is demanded that these be semantically null, even as the elements of the

phonemic or, for sign language, chereemic, alphabet are semantically null. Unless some considerations (such as those of Fodor 1966) are specified to justify the levels of analysis proposed, any behavioral repertoire can be decomposed to a semantically null level; in the present case, for example,  $\omega_i \alpha_j$  might be supposed to branch to the particular, semantically null, muscular sequence associated with it. Thus the oft-repeated<sup>25</sup> claim that the honey bee's dance language fails to meet this linguistic desideratum must be grounded in an approach along the lines sketched.

The underlying notion in Fodor 1966, on the other hand, seems to be that the demonstration of hierarchical structure is evidence for the inadequacy of a finite state account--evidence, that is, for the introduction of context-free phrase structure rules. Indeed, that discussion goes further, arguing that it supports a transformational analysis; "saying that bracketing of behavior is preserved under psychological operations that satisfy such formal descriptions as deletion, adjunction, permutation, etc. . . . is equivalent to saying that the grammar that underlies the behavior is transformational" (Fodor 1966:293). This is all very premature. The conclusion that the grammar is thereby shown to be transformational only follows if the "psychological operations" are grammatical as well as psychological. Finally, the demonstration

of hierarchical structure is not conclusive proof of context-free complexity. The tree-structure of (63), for example, is equivalent to the labelled bracketing of (64); strings of the form (64) are as accessible to a finite


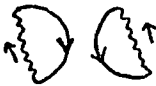
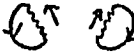
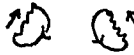


$$(64) \quad [Y \dots [W \overset{W}{i} \overset{\alpha}{j} ]_W \dots ]_Y$$

state treatment as was the unbracketed version: simply add the bracket symbols  $[Y, ]_Y, [W, ]_W$  to the terminal vocabulary and generate in the appropriate places. The context-free character of the base for human language is guaranteed by the observational adequacy argument outlined at the end of sec. 4.1.1--the base for the dance language remains finite state in nature. Any evidence of constituent structure or syntactic categories, from the wagging node  $W$  to a dance node  $D$  or dance-type node  $Y$ , can be accommodated in a purely finite state fashion.

Transformations that copy, delete, insert, move lexical material as do those in vogue for human language are only slightly tempting for the dance language.

As stated, sec. 3's mirror-image notes 3 and 6 are a part of the description of how to proceed from the sample paths of sec. 3 to the full finite state device. The paths given in the text consistently generate the left before the right circuit--the reverse generation is to be handled by the paths added by the augmenting procedures

of notes 3 and 6. What was viewed at that time as a space-saving device now merits consideration as an expression of cognitive space-saving by the dancer, i.e. as a

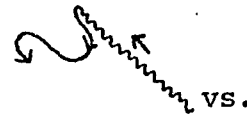
metathesizing transformation co-deriving  and ,  and ,  and .

The formalization of the proposed rule depends on a number of issues which have not been fully resolved, among them the syntactic status and nature of SD. In the event that it is SD and not D which is properly taken to be the syntactic domain, the degree of "sloppiness" tolerated by SD becomes a matter of some importance: how fixed is the L-R alternation of circuits in circuited SD's? If it is absolute, with  $(LR)^x$  and  $(RL)^x$ , integral  $x \geq 1$ , the only possibilities, then a metathesis account will have to be a metathesis "smear" rule, much as nasalization rules are in phonology; the optional metathesizing must be "smeared" throughout the input string whenever that option is exercised. If, on the other hand, L-R alternation in SD is not grammatically prized (or, for that matter, if it is D and not SD which is to be generated), the proposed metathesis is rather more analogous to a passivizing process, and has no such "smear" condition on

its application.

There are at least two alternatives to the transformational analysis response to what have here been termed mirror-image dance pairs. For one thing, there is some question as to whether the co-derived dances are in fact variants of the same message for a given dancer. In the event that dancers display a fairly rigid lateral preference in production and comprehension, allowing dancer-followers to be characterized as L-first or R-first circuiters, this transformational analysis would lose much of its plausibility. In the event of demonstrated lateral preference on the part of individual dancers, it is perhaps most naturally treated as yet another dialect discriminator, an approach which would be further supported if the dancer-follower were to demonstrate an equally strong

preference for one of the S-dance types (

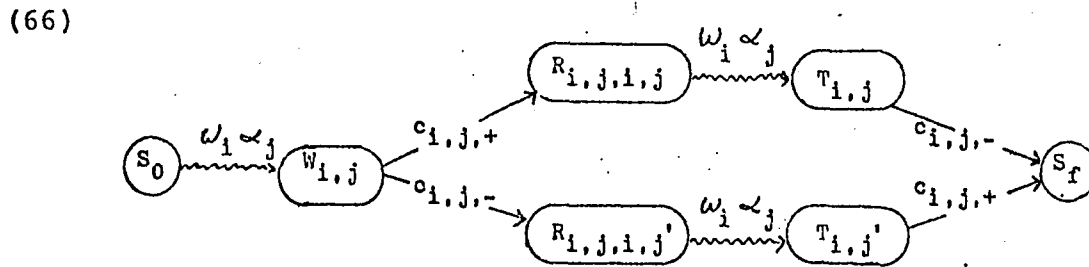


).  
 Then again, even if it should be the case that whatever can be a second circuit can be a first circuit for a given dancer, it is not immediately obvious that this would be a linguistically significant generalization, best captured by transformational rule. For in that case, there is no performance evidence for a "left circuit" as opposed to a "right circuit" category, no interesting

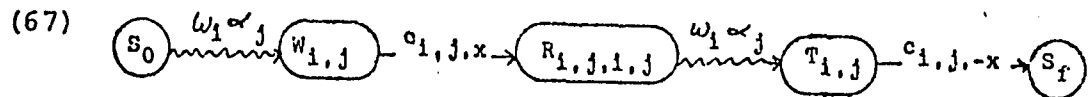
behavioral property attached to  $\{\omega_i \alpha_j \ c_{i,j,+} / 1 < I_1 \leq i \leq I_2 < m, 1 \leq j \leq n\}$ <sup>26</sup> as opposed to  $\{\omega_i \alpha_j \ c_{i,j,-} / 1 < I_1 \leq i \leq I_2 < m, 1 \leq j \leq n\}$ , or vice versa. In that event, circuits should not be described as L/R (or, to return to the notation of sec. 3, curved runs should not be described as +/-) but only as identical or opposite in orientation to antecedently or consequently executed circuits (curved runs) within D or SD. Following (65), the tailwagging dance which in sec. 3 required (see

- (65)  $x \in \{+, -\}$
- x = + if and only if x = -
- x = - if and only if x = +

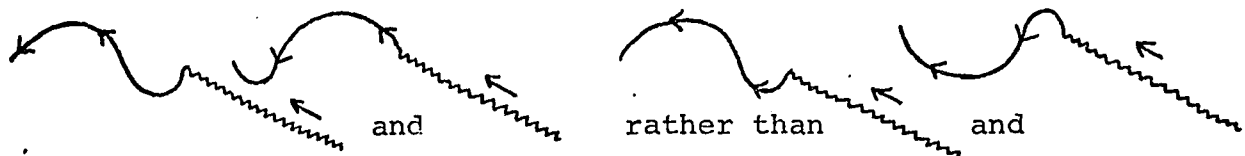
(12) and sec. 3 n.3) the seven-state, two path D- derivation of (66), reduced to the five-state one-path D-



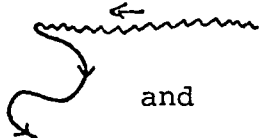
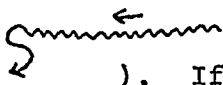
derivation of (67). The terminal vocabulary is in terms of x and -x, not the more specific + and - of sec. 3.



A transformational treatment must be motivated on three counts, then, with behavioral evidence to support a L vs. R/+ vs. - distinction at the level of the terminal vocabulary, evidence that it is an option for all members of the dance community, and finally, evidence that it is an option which is without semantic weight. Evidence for L vs. R rather than + vs. - is evidence that



are to be co-derived.<sup>27</sup> It remains something of a problem to decide whether the underlying form is to be in terms of L or R (or alternatively, + or -), and to motivate derivational ties between certain pairs of synonymous S-dances

and not others (such as  and ). If, however, all three criteria are met, some grammatical treatment is indicated, although not necessarily a transformational one; the series (68) of context-sensitive<sup>28</sup> rules accomplishes L-R metathesis, even if by introducing as yet

(68)      LR → AR  
           AR → AB  
           AB → RB  
           RB → RL

unsupported analytical stages (in the formulation given in (68), these are AR, RB, AB). Evidence for L-R metathesis with no evidence for intermediate stages, is evidence

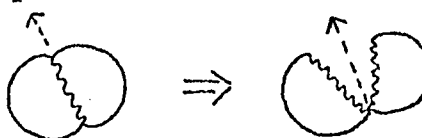
for transformational metathesis.

Typical human language operations of context-free constituent analysis, overlaid by a series of semantically inert transformations to delete, insert, move lexical material as per their context-free categorization are thus of quite limited usefulness in capturing dance linguistic patterning. Lexical dance language categorization is almost certainly finite state and the only transformational candidates considered to date (solar incrementing across SD, L/R or +/- metathesizing) are far from being well-supported. The search for the linguistically significant generalizations in the dance language may well be best pursued independently of the grammatical model for human language.

Supposing SD to be the syntactic domain, what is most striking about the dance language is the  $\omega$ -dependency of virtually all the phenomena. Thus, despite the previous discussion, it doesn't seem nearly as important to characterize all S-dances as S-dances, all tailwagging dances as tailwagging dances, etc. (which can be done by introducing such bracketings as  $[_S \ ]_S$ ,  $[_{TW} \ ]_{TW}$ , as was discussed with regard to (64)), as it does to somehow capture the correlation between each dance type and its distance range. From the point of view of the finite state device it can only be an accident that, for example, whenever an  $\omega_i$  value is shared by two dance types is is at

the extremes of the ranges appropriate to each, whenever  $\omega_i$  and  $\omega_{i \oplus p}$  ( $p > 0$ ) are appropriate to a given dance type, so is  $\omega_{i \oplus 1}$ ;  $\omega_{i \oplus 2}$ , ...,  $\omega_{i \oplus p}$ , whenever  $\omega_i \alpha_j$  is a semantic interpretation appropriate for a given dance type, so is  $\omega_i \alpha_1$ ,  $\omega_i \alpha_2$ , ...,  $\omega_i \alpha_n$ . The  $\alpha$ -dependency of the transitional dance patterning also seems quite significant and, unlike  $\omega$ -patterning, yields in an obvious way to a transformational treatment, which moreover unifies the full-wagging transitional and tailwagging dance types.

The proposal is quite simply to extend the  $\omega$  values for the tailwagging, i.e. (12), portion of the finite state base component to embrace all those  $\omega$ -values ultimately appropriate either to full-wagging transitional or tailwagging realization, deleting all paths of the full-transitional form (14). All full-wagging transitional dance forms are transformationally derived from underlying structure tailwagging in design; one of a series of "spread" transformations will pry the left and right circuits an equal,  $\omega$ -dependent number of  $\alpha$ -values away from the bisector which may be read off the underlying

tailwagging structure. So: 

The  $60^\circ$  limit on the angular deviation in such structures, which translated to the  $k = n/6$  (see (13)) limitation (14) is here reflected in the number of circuit-prying transformations needed. If L and R may differ from

each other by at most  $60^\circ$ , or  $k$   $\alpha$ -units, then they each differ from their bisector by at most  $30^\circ$ , predicting that  $K$  (see (69)) "spread" transformations will be required,

$$(69) \quad K = k/2 = n/12$$

prying the two circuits  $1, 2, \dots, K$   $\alpha$ -units away from their bisector. Since  $K$  must be an integer, this again establishes  $n$  as a multiple of 12.

Each transformation must be indexed for the  $\omega$ -values it ranges over; although most will be obligatory rules, it is conceivable that the transformation(s) having the least effect may be optional--in other words, that there are  $\omega$ -values for which minor deviation in the orientation of the wagging run is possible though not required (i.e. the  $\omega$ -ranges for the tailwagging and full-wagging transitional dance types overlap to some extent).  $\omega$ -range overlap for differing degrees of deviation from the bisector is reflected by overlapping in the  $\omega$ -ranges assigned the transformations productive of the degrees of deviation in question: (70) gives the "spread" transformation

(70) SPREAD1

$$\begin{array}{ccccccc} \omega_i \alpha_j & c_{i,j,x} & \omega_i \alpha_j & c_{i,j,-x} & \text{optional?} & & \\ \omega_i \alpha_{j \ominus 1} & c_{i,j \ominus 1,x} & \omega_i \alpha_{j \oplus 1} & c_{i,j \oplus 1,-x} & \implies & & \end{array}$$


Conditions:  $x \in \{+, -\}$ ;  $-x = +$  if  $x = -$ ,  $-x = -$  if  $x = +$ ;  $1 \leq j \leq n$ ,  $1 < I_1 \leq i \leq I_2 < m$ ,  
where  $I_1$  and  $I_2$  are the constants defining the  $\omega$ -range appropriate to this degree of angular deviation.



appropriate to the largest  $\omega$ -values short of actual tail-wagging dances. As before (see (40)),  $\ominus$  denotes the arithmetic inverse of  $\oplus$ . In (70),  $x$  is a variable even as  $i$  and  $j$  are: its status is very different in (65), where it functions as a lexical symbol on the order of  $\omega_1 \propto 6$ .

The evidence for a variable  $x$  in the structural description of spread transformations is indirect: without such a variable, evidence for spread transformations is evidence for a later metathesis transformation (or



$$\left\{ \begin{array}{c} + \\ - \end{array} \right\} \Rightarrow \left\{ \begin{array}{c} - \\ + \end{array} \right\}$$

transformation). For, without a transforma-

tion to provide a unified treatment of  and

, the spreading of  requires

two rules: one spreading  to , and

the other spreading the corresponding  to 

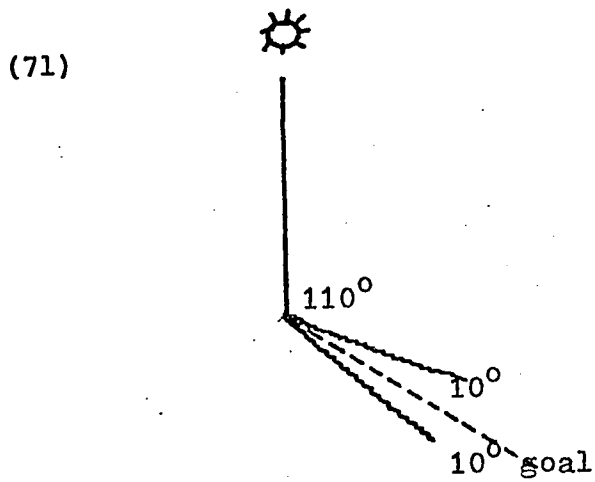
(again, presumably). In other words, without a variable  $x$ , full-wagging transitional dances are derived either by  $K$  spread rules followed by a metathesis (or +/- switching) rule or by  $2K$  ( $=k$ ) spread rules. Not only does the metathesizing (+/-) analysis feature fewer ( $K+1$  vs.  $2K$ , with  $K$  large since  $K = n/12$ ) transformational rules, but of course the finite state device is also thereby simplified.<sup>29</sup>

Without a variable  $x$ , evidence for spread rules is con-  
founded with evidence for a metathesis (+/- switching)  
rule.

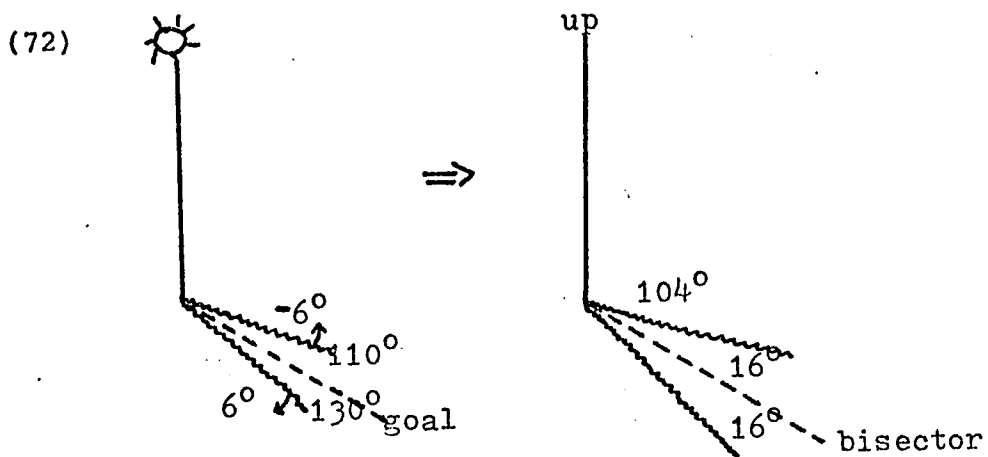
Again, the evidence with regard to spread rules is  
at best equivocal. On the positive side, the peculiar  
distribution of wagging pairs in (14) (i.e.,  $\omega_i \alpha_j$  pairs  
with  $\omega_i \alpha_{j \oplus 2}$ ,  $\omega_i \alpha_{j \oplus 4}$  etc.--see p. 115) is "explained  
by the integral spread rules (SPREAD 1, SPREAD 2, ...,  
SPREAD K) whose  $\omega$ -value indexing makes explicit the pat-  
tern which the finite state account could only hint at by  
way of brackets of the form [<sub>SPREAD1</sub> ]<sub>SPREAD1</sub>. The  
fact that  $\omega$ -values found within a given pair of spread  
brackets fall into a certain range is elusive, a spe-  
cial case of the complaint made earlier concerning the  
correlation between each dance type and its  $\omega$ -range. On  
the negative side, evidence for at least some versions of  
the SD claim is, at least to some extent, evidence against  
spread rules, for two reasons: (a) the "average to an  
integral  $\alpha$ -value" constraint will be so much more ac-  
cessible to a semantic rule of averaging defining anomaly  
as non-integral output, than to a syntactic transforma-  
tion or series of transformations designed to skew the  
 $\alpha$ -values in an integral-neat but random,  $\omega$ -dependent way;  
and (b) the "average to an integral  $\omega$ -value" constraint  
appears to have roughly the same cognitive, linguistic  
status and is therefore probably best accorded parallel

treatment and yet there is no reason whatever for supposing this to be a syntactic phenomenon, no dance type which varies  $\omega$  in a systematic way the way transitional dances vary  $\alpha$  so that more than one wagging run is needed to determine that parameter of the message.

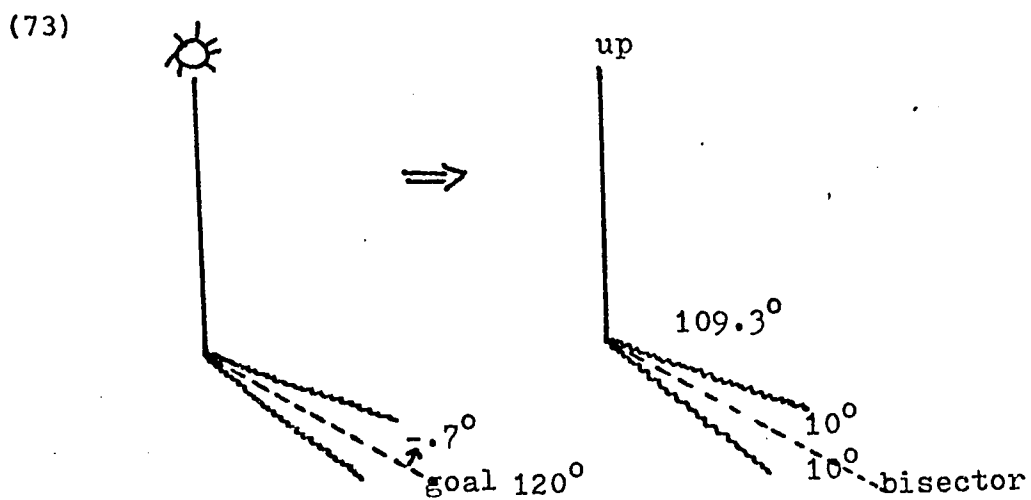
Following Fodor 1966, it can be asked what behavioral evidence there might be for a two-level analysis of the transitional dance. A natural response would be in terms, once again, of residual misdirection data, the dance form of interest, the transitional dance. Residual misdirection "disappears when the earth's magnetic field is artificially eliminated" (Gould 1976:215, reporting results of Lindauer and Martin 1968) and thus is the requisite physical performance character. Since direction may be given by runs as much as  $60^\circ$  apart and the quality (clockwise, counterclockwise) and extent of the misdirection depends in a finer-than- $60^\circ$  way upon the run's relationship to the horizontal and vertical axes (see (7)), it should be possible to test statistically whether, in transitionals, it is the executed angularly deviating runs or the non-executed bisector of those runs which is being misdirected in this systematic way. To take an example, consider (averages of) a dance which on a horizontal dance floor is executed as in (71); it is subject to residual misdirection (following (7)) on the vertical dance floor BUT is it the two executed wagging runs which are



affected, as per (72), or the bisector which is affected,



as per (73)? According to (7), although for many transitional



dance  $\alpha$ -pairs<sup>30</sup> the two hypothesis give the same results, yet there are also many  $\alpha$ -pairs<sup>31</sup> for which, like the 110° and 130° pair discussed here, the two hypotheses differ. Specifically, these differ with regard to whether the dependence of the angular deviation in transitional dances on  $\omega$  is constant, as in (73); in (72), on the other hand, their relationship depends upon the physical setting of the dance, with residual misdirection perturbing the correlation. If, that is, it is the executed runs which are geometotactically misdirected as in (72); then a specific deviation  $d$  would hold for  $\omega_d$  on a horizontal surface for all bisector  $\alpha$ -values, but only for certain  $\alpha$ -values (see n.30) at  $\omega_d$  on the vertical surface. Finally, for each other  $\alpha$ -value such as  $\alpha_x$  and  $\alpha_y$  (see n.31) there will be a  $\omega_{d_x}$ ,  $\omega_{d_y}$  (different from  $\omega_d$ ) such that  $d$  is the correlated deviation for that angle on the vertical surface (though again not on the horizontal surface or for other  $\alpha$ -values on the vertical).

In sum, the misdirect-the-wagging runs hypothesis deals with the physically observable linguistic signal but complicates the statement of the dependency of angular deviation on  $\omega$ .<sup>32</sup> The misdirect-the-bisector hypothesis, which it has been suggested would provide performance support for a series of spread transformations of the indicated kind, leaves the angular deviation dependent solely on  $\omega$ . Interestingly, von Frisch's computation of the

individual residual misdirection values, reported in (7), assumes the misdirect-the-bisector hypothesis.

In each case, the value was computed (see von Frisch 1967a:214) by averaging all wagging runs (whether produced in the course of an S-, tailwagging or transitional dance) advertising a goal situated at the specified sun-relative angle with respect to the hive. This means that if, for example, a goal at an angle of  $130^\circ$  from the sun provokes a transitional dance whose wagging runs give the distance as  $\omega_i$  and the direction as, say  $140^\circ$  and  $132^\circ$ , it makes a  $6^\circ$  contribution  $((140^\circ - 130^\circ) + (132^\circ - 130^\circ))/2$  for  $130^\circ$ , which of course is identical to the contribution that results from calculating the deviation of the bisector of  $136^\circ$  (i.e.,  $(140^\circ + 132^\circ)/2$ ) from  $130^\circ$ .

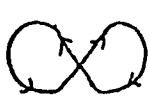
This assumption of his has the consequence that (7) can not be used (as it was above for purposes of illustration) to delineate the angular values which discriminate the two hypotheses; residual misdirection values must be computed solely on the basis of S- and tailwagging dance data, for which dance types the two hypotheses make identical claims. The residual misdirection table so computed then provides the standard against which transitional dance residual misdirection can be investigated, as indicated.

It would be very satisfying at this time to be able to conclude that the discovery that the bisector is residually misdirected in transitional dance forms would be

conclusive proof of transformational grammatical complexity. There remain at least two alternative explanations for such an eventuality, however. Given the concept of the dance (D) outlined in these pages, such a discovery does demonstrate that there are two analytical levels to be represented in the grammar for transitional dances. What it does not demonstrate is that these levels are both syntactic in nature.

This is an issue which is virtually uniformly ignored in the-power-of-the-grammar literature, although it often in fact plays a crucial role. In the present case, it is by no means clear that a one-component grammar with syntactic spread transformational rules is to be preferred to a one-level finite state bracketing syntax in tandem with a series of semantic rules averaging the divergent transitional waggling runs to a simple waggling-run semantic representation.

Doubt would be cast on the one-component grammar with spread transformations analysis by any of the following: (a) the discovery that in dances of the forms



and



(which can not be generated from

another dance form with overlapping runs by an optional, or even  $\omega$ -sensitive obligatory, spread rule), it is the bisector which is subject to residual misdirection; (b) performance evidence that the syntactic domain is SD, with a

high degree of "sloppiness" which is not merely tolerated but further, is to be utilized in determining the message, so that the effect of geomenotactically misdirecting the bisector on a left or right circuit depends on the ratio of left to right circuits in that SD. In either case, simplicity considerations would dictate a solution in terms of an averaging semantic component, perhaps (at least for (b)) with semantic as well as syntactic well-formedness constraints, stigmatizing configurations which do not average integrally as semantically anomalous.

What a residually misdirected bisector establishes therefore is that the message level is linguistic and not merely a matter of interpreting the formal system, outside the formal system itself. Within the context of the discussion here, this can be accounted for either by a syntax of transformational complexity, or a two-component grammar uniting a finite-state categorizing syntax with an averaging semantics perhaps introducing a "semantically anomalous" tag onto certain syntactically well-formed strings. These are the hypotheses which integrate most neatly with the grammatical discussions of sec. 3 and sec. 4.1.1. Yet a third explanation surfaces upon examining the informal linguistic reasoning behind von Frisch's assumption of bisector misdirection in (7).

Von Frisch's misdirect-the-bisector bias does not stem from any predilection for spread rules, but rather from

his assumption that all waggling runs, regardless of dance type, are direct attempts at representing the angle between the goal's and solar azimuths. In linguistic parlance, von Frisch treats the angular deviation in transitional dances as a performance problem, i.e., an execution error. What's more, it is the waggling runs' deviation from the "expected" angle (that between the azimuths of the goal and the sun) and not their deviation from each other or their bisector which absorbs his attention. This (implicit) linguistic analysis is almost universally accepted in the literature. This analysis also succeeds in unifying the full-waggling transitional and tailwagging dance types and, it might be argued, at much less cost to the grammar than, for example, either syntactic spread rules or a semantic average-to-an-integer constraint; transitional dances are simply error-ridden tailwagging dances.

There is no experimental test which will discriminate between this position and that taken earlier, viz. that the transitional dance is, at one level of grammatical analysis at any rate, different from the tailwagging dance type. Not even a demonstration that, for example, the bisector of horizontal surface transitional dances is statistically precisely as accurate an angular measure as are the coincident waggling runs of the tailwagging dance, can preclude the von Frisch "error" analysis; that analysis

could accommodate such an experimental observation by supposing that there is a two-stage performance interference at small (i.e., transitional)  $\omega$ -values, the "spreading" performance factor being peculiar to those values. And if it proved to be the executed wagging runs which are residually misdirected, that too could be accommodated by a two-stage performance interference at small  $\omega$ -values, the "spreading" peculiar to the particular  $\omega_i$  preceding the more general residual misdirection perturbation.

Granted that neither hypothesis is experimentally falsifiable, it may yet be instructive to examine the theoretical underpinnings of the two approaches. For the biologist, the dance language is interesting as an encapsulating report of the dancer's recent successful search behavior; insofar as that report deviates from a strict channel-switching repetition of the behavior, it is naturally stigmatized as an "error." For the linguist, however, the dance language is interesting quite apart from its reporting function, in fact principally not for its reporting function at all but rather for its organizational properties; in a real sense, the dance language is interesting precisely insofar as it deviates from a strict channel-switching repetition of the behavior. What's more, if the grammatical analysis of the dance language is to be compared with that of human language, it is important to give the honey bee grammar the benefit of the


doubt, so to speak, so as to ensure that species chauvinism is not the key factor accounting for the results of that comparison.

For the biologist working within a behaviorist framework, the linguistic reality of the stimulus input (realized here as the bisector) needs no demonstration; what requires demonstration is that the organism makes some positive linguistic contribution to the signal apart from a threshold perceptual categorizing. Hence, the transitional dance is an error. For the linguist working within a rationalist framework, all abstractions from the physical signal must be justified, including the most superficial linguistic representation of that signal, in terms of the cognitive apparatus behaviorally evinced. Which leaves us with the problem of linguistically motivating a decision between these two views of the data.

The von Frisch hypothesis is to some extent linguistically vulnerable due to its ad hoc character. It is, after all, solely this unsupported channel-switching premise which dictates that the transitional dance (and not the S-dance or round dance or ...) is an error-ridden tailwagging dance. Again, supposing there to be an angular-error-causing factor operative only at small  $\omega$ -values, there is still the problem of explaining why this error is leftward on left circuits, rightward on right circuits, leaving their bisector a reasonably good




directional indicator. Why, in other words, doesn't this execution handicap destroy the message? In the circumstances, it seems preferable to suppose that there are two message-coding dance language conventions; each restricted (for reasons as yet unknown) to a range of  $\omega$ -values.

#### 4.1.3 "Explanatory" Adequacy

In the course of investigating the observational adequacy of sec. 3's finite state grammar, two central issues emerged: (a) the constituency of SD, assuming SD to be the syntactic domain of the dance language; and (b) the description and grammatical status of solar incrementing. The generation of SD as a series of D's is easiest to accommodate; the no-D version of the SD claim is expedited by an averaging semantic component identifying the meaningful grammatical utterances with those whose  $\alpha$  and  $\omega$  subscripts each average to an integer. Solar incrementing of the danced angle  $\alpha$  during a SD, if syntactic, is most likely transformational; the alternatives explored are pragmatic, one straightforwardly incrementing the angle, the other translating the variable direction with respect to  to a constant direction with respect to a hypothesized visual or olfactorily obtrusive landmark A in the hive vicinity, assumed to be a key component in the forager's initial perception of solar

azimuthal change. Experimental evidence, albeit largely anecdotal, is presented which argues against a syntactic treatment of this phenomena by indicating incrementing ability in nonlinguistic tasks. Sec. 4.1.1 concluded with an outline of the observational adequacy arguments that have been mounted against a finite state or context free syntax for human language: the structural claims which parallel argumentation for the dance language case requires are at best untenable at the present time. The verdict: observational adequacy may well be assured by a finite state treatment not very far removed from sec. 3's account.

In sec. 4.1.2, the issue of descriptive adequacy was addressed, principally with regard to determining what grammatical (competence) constructs are confirmed as valuable input to a performance theory of the simplest kind. Since to a great extent, the relevant experimental work has not been done, the discussion was most often in the form of what would constitute behavioral (performance) evidence for one grammatical analysis rather than another. Transformational syntactic complexity would be indicated by a demonstration that only in full-wagging transitional

dances (i.e.,  , but not  or  ) is it the bisector which is residually misdirected--in all other cases, it is the executed runs which are deformed.

Uniform residual misdirection of the bisector suggests a two-component grammar with a finite-state syntax and an averaging semantics, perhaps even a three-way division of dance maneuver concatenations: ungrammatical, grammatical and meaningful, grammatical and semantically anomalous.

Chomsky 1964:28 identifies three levels of adequacy to be met by the correct grammatical description for a given human language. Observational adequacy "is achieved if the grammar presents the observed primary data correctly." Descriptive adequacy "is achieved when the grammar gives a correct account of the linguistic intuition of the native speaker, and specifies the observed data (in particular) in terms of significant generalizations that express underlying regularities in the language." The investigation of the dance language of the honey bee within these terms, i.e., observational and descriptive adequacy, is particularly problematic; the commitment to an infinite set of grammatical dances which is presupposed by the descriptive adequacy issue is, perhaps, not immediately obvious or particularly palatable. Explanatory adequacy, however, is irretrievably acquisition-based, and, as such, apparently does not apply to the honey bee, whose linguistic system is not acquired but inherited. (74) = (Chomsky 1964's 1b) is central to the three-way distinction drawn:

(74) primary linguistic data → **B** → generative grammar

a grammar that aims for observational adequacy is concerned merely to give an account of the primary data (e.g., the corpus) that is the input to the acquisition device (lb); a grammar that aims for descriptive adequacy is concerned to give a correct account of the linguistic intuition of the native speaker; in other words, it is concerned with the output of the device (lb); and a linguistic theory that aims for explanatory adequacy is concerned with the internal structure of the device (lb); that is, it aims to provide a principled basis, independent of any particular language, for the selection of the descriptively adequate grammar of each language. (Chomsky 1964:129)

The goal of linguistics, then, is a theory which "offers an explanation for the intuition of the native speaker" (Chomsky 1965:25-26). Or, again, "[a] theory of linguistic structure that aims for explanatory adequacy incorporates an account of linguistic universals, and it attributes tacit knowledge of these universals to the child. . . . Language learning would be impossible unless this were the case" (Chomsky 1965:27).

Granted that it is acquisition which is in most need of explanation in human language, it is clear that this is not by any means helpful as regards the task of justifying an analysis of the dance language. It does not follow from this, however, that dance language grammatical evaluation must be pursued solely on observational and descriptive grounds. The notion of explanation relevant to a genetically coded system is teleological "explanation." What is needed is a dance language grammar which makes some evolutionary sense of the system itself,

providing some "explanation" for the uniqueness of the directional signal as compared with the multiplicity of distance signals, the syntactic variety in dance types and their essential preservation across races of honey bee, with each race having roughly the same number of dance forms, with only the transitional dance shapes varying across dialects. There is also the matter of accounting for the interracial agreement that the round dance be restricted to nearby goals, the S-dance to distant goals, that angular deviation across SD vary inversely with distance. Why?

To paraphrase Chomsky, "explanatory" adequacy is achieved if the grammar is a teleologically principled descriptively adequate account. It must be emphasized that it is not the adaptive value of the dance language which is relevant here.

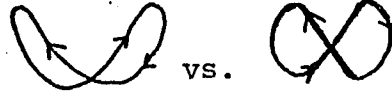
The dance language controversy, which was outlined in sec. 2.13, was initially a confrontation between a naive functionalism (see, for example, Lindauer 1971 and von Frisch's 1968 "Do bees really not understand their own language?") and a demand for empirical justification of that functional account (see, for example, Wenner 1967). The outcome of that controversy was a highly refined experimental paradigm (Gould 1975b) allowing for an absolute separation of linguistic from nonlinguistic communication for this species, demonstrating the

communicative efficacy of the dance language. The injection of functionalism into the linguistic analysis may, given the recent dance language controversy experience with teleological considerations (Ferguson 1975 notwithstanding), be viewed as suspect. However, teleological considerations remain, along with a simplicity criterion, the sole guides to grammatical hypothesis formulation, with the hypotheses so formulated required to be susceptible to empirical testing (on observational and descriptive grounds, at least), even as Wenner proposed to test that of von Frisch, thus creating the dance language controversy. The search for a functional grammar is not at all dependent on the identification of the adaptive value of the dance language itself. A grammar achieves "explanatory" adequacy if the structural claims, demonstrably descriptively adequate, are teleologically helpful, offering some hope of answering the heritability issues addressed in the previous paragraph. The cognitively based digital account of secs. 3-4.1.2, although arguably descriptively adequate is apparently not adequate in this respect. The question is whether the implicit linguistic bedrock of current experimental work is in any way indicative of a solution.

Apart from vague and rather naive flirtations with functionalism, there has been very little evidence in the literature of a concern with teleology, at least not as it

relates to structural aspects of the dance language. To be sure, von Frisch (1967a:297) does express concern for the genetic packaging of the dance language dialect fea-

tures of distance coding and



transitional dance type, specifically its relationship to the genetic packaging of racial color-marking (see sec. 2 n.11). And Gould (1975a:692) speculates in a footnote on the adaptive role of the dialect-dependent error magnitudes as follows:

One possibility might be that the reduced angular scatter compensates for increasing distance; and that, taken with the nearly constant distance scatter, could serve to generate the same recruit distribution about a target regardless of its distance from the hive. . . . If food sources exist in "patches," it could be to a colony's advantage to spread out its searching recruits to exploit the patches more efficiently. The area around the hive for which round dances are performed might be thought of as the first such "patch." Italian bees have both a smaller round dance zone and recruit scatter than Carniolan bees. If this difference has (or had) an adaptive value, one might suppose that the two races evolved in habitats with different patch sizes or colony spacings, or that some compensating difference in search strategy in the field may exist.

This remark, although apparently in keeping with his demonstration of linguistic communication, is strikingly inappropriate given his repeated assertion<sup>33</sup> that recruitment is in all likelihood ordinarily achieved by non-linguistic olfactory cueing--accounting for the experimental results of Wenner and others. Only behavior with a highly significant recruiting role would be subject to

selective pressures of the kind outlined; if olfaction plays a large part in recruitment, down to the final (goal-finding) stage in dancing-instigated search behavior,<sup>34</sup> then the size of the typical feeding site's "profitability radius" clearly could not exert an influence on the error quotient of the dance.

In large part, the earlier issue (see sec. 2.4.2) of distance vs. flight range percentage as the semantic weight of the  $\omega$ -variable was in response to the teleological question raised by von Frisch's use of the term "dialects." Gould's teleological query is expressed above in terms of a correlation between SD angular deviation and wagging duration. Neither conception of the problem materially contributes to its resolution, which persists as perhaps the most serious "explanatory" grammatical hurdle.

Dethier 1957 is a report of "dance" behavior in the fly *Phormia regina*. The dance elements which recur in this species, to what purpose (if any) is not known, are three: the behavior is released by a taste stimulus of a certain threshold which is dependent on the fly's pre-stimulus hunger-state, the circular turning "searching" movements thereby elicited are vaguely dance-like, and the patterns are relatively symmetrical about an introduced light-beam or gravity axis. Although this clearly represents an attempt to widen the biological perspective so as to shed some light on the teleological issue, it

remains of only peripheral interest<sup>35</sup> here, given the very small number of linguistic resemblances which can be noted. From a grammatical, competence point of view, the comparison of the fly and the honey bee sheds about as much light on the structure of the dance language as the comparison of man and the "Polly wants a cracker" parrot does on the structure of English.

The general approach is not necessarily a poor one, but what is needed for a useful linguistic comparison is a species displaying an inherited system at least as close to the bee's as the acquired songbird's is to man's.<sup>36</sup>; teleological resemblances in the first case, acquisition resemblances in the second are what it is crucial to uncover. Dethier, as far as can be told from his discussion of the case, simply chose the wrong species.

Pankova 1972 takes a very different tack, proposing, in effect, an alternative grammatical analysis to that of von Frisch (as sketched on pp. 188- 89) and, for that matter, that pursued here. The overriding consideration again is teleological, specifically the explanatory underpinnings of the syntactic variation in dance type. Unfortunately, the Pankova account, in neatly identifying dances of different types, manages to preclude a differentiation at any more superficial structural level. As such, it suffers from the same problems, though as will be argued to a much greater extent, as von Frisch's

channel-switching syntactic analysis did. That analysis unified full-wagging transitions and tailwagging dances by labelling the angular deviation in transitionals as an execution error; Pankova's analysis manages to present a unified account of the round and tailwagging dance forms (and presumably transitionals as well, although these are not discussed as such), by giving the behavior a structural description even further removed from the physical display coding the message.

The conclusion, that "the structural elements of both the wagtail and round dance are in all probability identical and the signal system in honey bees is unique" (Pankova 1972:407) depends upon a structural description based not on  $\alpha$  and  $\omega$  features or even wagging vs. curved runs, but rather the variation through time of the longitudinal axis of the dancer's body. The observation that "the trajectories of . . . the middle of the parietal region, the center of the thorax and the abdominal end . . . described during the dance an alternatively curved motion and their speeds and accelerations were different" (Pankova 1972:405) motivates the retreat to the more abstract (?) structural level represented by the longitudinal axis. The subsequent demonstration that assigning angular values to the different longitudinal axis positions produces a (roughly) sinusoidal curve over time for either dance type accounts for the conclusion drawn

above.

The sole reported difference in the curves is in the duration of the shift in rectilinear motion, which of course is the correlation with distance. The implication seems to be that Pankova's honey bee, Apis mellifera caucasica Gorb., may hint at the direction toward the goal during the round dance by clustering the reversals in that direction (in the terminology developed earlier, by interpolating sickles in an otherwise round SD); it is difficult to be sure, but that seems to be the sense of the remark that "[t]he round dance does not include information about the distance only" (Pankova 1972:407).

In a real sense, the precise machinations of an analysis structurally unifying the tailwagging and (true) round dances are uninteresting. The central, inescapable fact from the linguistic point of view is that the first but not the second gives distance and directional information. Explanatory adequacy is not achieved by ignoring that fact.

The problem of developing a dance language grammar which is "explanatorily" adequate remains largely unsolved. What is necessary here, as elsewhere in linguistics, is "to discover a complex of data that differentiates between conflicting conceptions of linguistic structure in that one of these conflicting theories can describe these data only by ad hoc means whereas the other can explain it

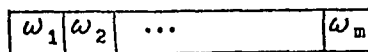
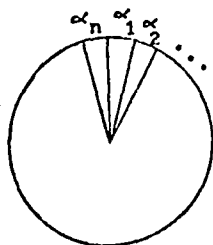
on the basis of some empirical assumption about the form of language" (Chomsky 1965:26).

#### 4.2 Lexical Criticisms of the Finite State Account

Thus far, the criticisms levelled at sec. 3's finite state grammar have all been structural in nature, concerned with the hypothesized rules for combining "lexical items" ( $\omega_i \alpha_j c_{i,j,-}$  etc.) into "sentences" (D or SD). Any grammar, however, and sec. 3's is no exception, is vulnerable not only on structural but also on lexical grounds--vulnerable, that is, to criticisms directed against the nature of the hypothesized lowest level of analysis.

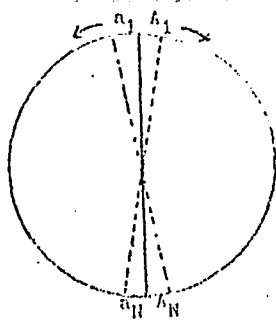
There are three lexical positions which have been taken in the literature, in each case with little or no argumentation. Sec. 3's lexical position, once again, for comparison, is that D (or SD) is a structurally constrained concatenation of maneuvers  $\omega_i \alpha_j, c_{i,j,+}, c_{i,j,-}, S_{i,j}, L_{li,j}, L_{ri,j}, R_{li,j}, R_{ri,j}$  where for some (so far undetermined) integral values  $m$  and  $n$  greater than 1,  $1 \leq i \leq m$  and  $1 \leq j \leq n$ , as depicted in (9), repeated here as (75) for convenience. The two grids

(75) = (9)



may be anchored at  $\omega_1 \alpha_1$ , the wagging run indicating the smallest danced distance value with the goal directly toward the sun. The different lexical positions taken are essentially directed at (75), and indeed generally at either the m-way division of  $\omega$ - or the n-way division of  $\alpha$ -values.

This account of the angular discrimination vocabulary differs from that presented in von Frisch 1967a:204, which argues for a 180° sweep of the left and right sides, as in (7), which "corresponds logically to a goal that is situated to the left or right of the sun." The proposal then is for a revision of the  $\alpha$ -variable as per (76)<sup>37</sup>. A 360° analysis such as (75) is explicitly rejected, for (76)

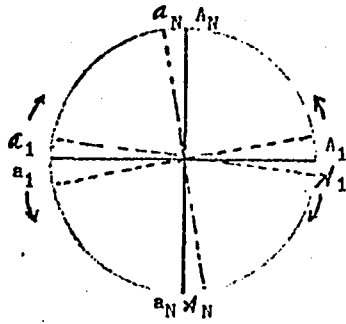


overly complicating the account of residual misdirection: "reckon[ing] the angle of dancing through 180° right and left from 0° for the direction vertically upward . . . yields a positive value for the misdirection [whether the

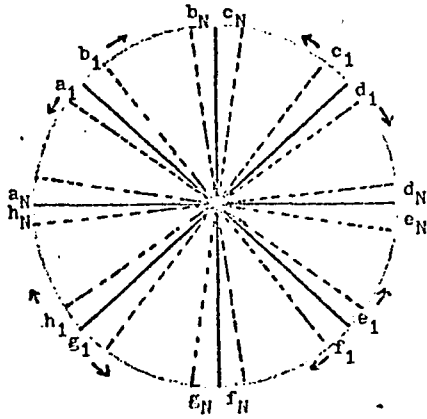
goal is  $150^\circ$  to the left or to the right of the sun]; with continuous reckoning over  $360^\circ$  (counterclockwise) there would be a positive misdirection on the left and a negative one on the right" (von Frisch 1967a: 205). This amounts to an embryonic competence/performance distinction on von Frisch's part, with the competence model serving as input to the production system, viewed as a noise-contributing filter. While the theoretical underpinnings of his analysis are, therefore, quite congenial, the particular analysis is not well-supported.

In the first place, a halving of the conceptual space along the vertical can not account for the residual misdirection deflections toward the horizontal axis, which is clearly of equal grammatical weight even if it is quantitatively less pronounced a phenomenon (von Frisch 1967a:216). In fact, it does not even provide the simplest input to a toward-the-vertical deflection operation, since it requires angular subtraction at small  $a_i, A_i$  values but angular addition at large  $a_i, A_i$  values (such as the  $150^\circ$  case of the previous paragraph). The  $\sphericalangle$ -grid in (77), which uniformly treats residual misdirection as an addition operation, thus presents a better analysis of vertical misdirection. At first blush, it seems that the inclusion of horizontal misdirection requires an analysis along the lines of (78) rather than (77). This is not the case, however, some lexical distinction must be

(77)



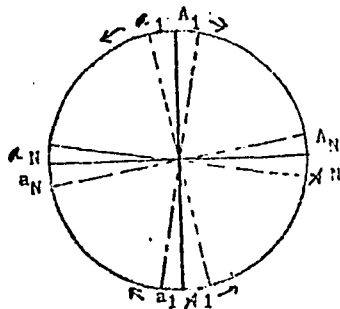
(78)



drawn between misdirectable-toward-the-horizontal and misdirectable-toward-the-vertical angular values, given that the magnitude of the misdirection varies, with the vertical exerting a much stronger attraction. This being the

case, (77) seems preferable to (78): small subtractions occur when the subscript of  $i > 1$  of  $\alpha_i \in \{a_i, A_i, a_i, A_i\}$  is small, while (comparatively) large additions occur when the subscript  $i$  is large. There can be no evidence for (77) in preference to (79), with large subtraction at low

(79)



$i$  values, slight addition at large  $i$  values, any more than there could have been evidence for a clockwise (see (75)) rather than a counterclockwise 360° display of the supposed  $\alpha_1, \alpha_2, \dots, \alpha_n$

The question is whether the goal of descriptive adequacy, the capturing of the linguistically significant generalizations of the dance language, is more easily met by (75) or by (77). Just how great a role do performance data play in this case?

The observation that an analysis along the lines of (76) or (77) has a certain "logical" significance (see p. 203) vis à vis the left vs. right positioning of the goal with respect to the solar azimuth has no apparent

weight. Any number of conceptual categorizings might be so justified--the force of the adjective "logical" appears to be that the proposed categorizing exhaust the possibilities. It has not, for example, been suggested that the angular "error" (deviation from the bisector<sup>38</sup>) in transitional dances is eye-dependent, with the legs of a to-the-left-of-the-sun goal restricted to left angular values (the  $a_i$  in (76)).

The evidence, then, for (77) rather than (75) is drawn solely from the patterning of the residual misdirection perturbations. So, once more, how relevant is this data to a theory of dance competence? Granted that distinctions proven relevant at the performance level must originate in the grammar, does it follow that the absence of a distinction at the performance level (as between the residual misdirection proper to the angular values of  $150^\circ$  to the left/right of the sun) is due to the absence of a (here, left/right) distinction in the grammar? Is the observed symmetry of a clearly physical, perceptual process, tied to the earth's magnetic field, to be supposed a reflection of an underlying conceptual symmetry? Put this way, it suddenly loses much of its plausibility. The notion of descriptive adequacy as preferring the simplest possible input to the performance system clearly requires some clarification; the very distinction between competence and performance is endangered

by too enthusiastic a commitment to input simplicity in this case.

It must be noted that (77) is not suspect because it misses capturing such linguistically significant generalizations as the  $j \oplus 1, j \oplus 2, \dots, j \oplus k, k = n/6$ , constraint on the right circuit wagging run of an  $\alpha_j$  left circuit wagging run transitional dance (see 14)). That is, the fact that the  $\alpha$ -values that would pair with, say, a transitional left circuit  $a_n$  are not  $a_1, a_2, \dots, a_k$  (see sec. 3 n.5), but rather  $A_1, A_2, \dots, A_k$  is not really problematic. No way has thus far been found of capturing this 60° limit; the  $k$  which figures in (14) has (hopefully) heuristic value for the reader of sec. 3, but certainly no linguistic status in the finite state account. If sec. 4.1.2's spread transformations (see (70)) were grammatically established, a case could be made for (75) rather than (77) on grounds of rule-counting simplicity: there would be fewer spread rules given (75) than given (77). As it is, however, there is no structural objection to (77).

The objection to (77) is rather that it considers physiological limitations within the competence account. A natural revision of the simplest-input-to-performance criterion for descriptive adequacy is one which, again, deduces competence distinctions from behaviorally attested performance distinctions. Simplicity of input, however, is apparently a desideratum only when the behavior

may be presumed to have some cognitive basis, which notion of course resists operational definition. To be sure, there is no evidence against (77) but neither, it seems, is anything unequivocally in its favor. The similar misdirection output in the two  $150^\circ$  cases may be explained either by a similar competence input, as by (77), or by an overlaid symmetrical magnetic field induced residual misdirection operation. The observed symmetry simply (no pun intended) seems more likely to be a contingent consequence of the biological structures mediating the gravity orientation process than direct evidence for a particular lexical structure. Other factors which are likely (and, admittedly, likelihood is in the eye of the beholder) to prove to be biologically interesting but linguistically irrelevant include those responsible for the directional indication in  $180^\circ$ -displacement trained transport cases (see p. 47) and the fact that the presence of polarized light leads to perceptual averaging whereas a view of the sun does not (see p. 71).

What is more to the point at this time is the overall impact on the lexicon of the proposed elimination of physiological limitation considerations from the grammatical account. For, the assumed finite number of angular discriminations is suddenly called into question. Both (75) and (77) make the unwarranted assumption that the conceptual space divides into a large but finite number.

of discriminable values, (77) additionally supposing the values to be arranged in quadrants. Discrimination ability must not be methodologically confused with conceptual categorizing.

To put the point somewhat differently, an  $n$ -ary  $\alpha$ -grid presumes that each  $\alpha_i$  has a certain width. Regardless of how large  $n$  is, and therefore how small the width of  $\alpha_i$ ,  $\alpha_i$  will be a cover term for a non-denumerably infinite number of directional values. In effect, the  $\alpha$ -grid defines  $n$  competence equivalence classes of possible danced directions. No evidence has been or can be offered for this rather substantial cognitive claim. The simpler, more conservative conceptual position is that there is no cognitive substructure of this kind with regard to directional indications. In other words, that the  $\alpha$  variable takes on, not a finite set of values  $\alpha_1, \alpha_2, \dots, \alpha_n$  as in (75), or a finite set of values  $a_1, a_1, A_1, \mathcal{A}_1, a_2, a_2, A_2, \mathcal{A}_2, \dots, a_n, a_n, A_n, \mathcal{A}_n$  as in (77), but a nondenumerably infinite number, each cognitively distinct, if not perceptually discriminable, from all its fellows. The immediate implications of this position are rather startling, since no formal system, not even one of transformational capacity, can generate a non-denumerable language.<sup>39</sup> The conclusion: the dance language is syntactically different in kind, rather than degree, from human language.

Gould 1975b:191 claims that both distance and directional dance linguistic parameters are analog: "In . . . graded signals, . . . the (analog) value being signalled varies about an average in a statistical fashion (as in the distance and direction indications of the dance)." Direct observation--"No two waggle runs seem the same either in distance or direction" (Gould 1975b:171)--cannot be decisive here for, regardless of the cardinality of the  $\alpha$  and  $\omega$  grammatical repertoires, it is to be expected that there is a noise-contributing performance filter operative in production and another in comprehension. It cannot be supposed that every minute but measurable distinction which can be identified in performance is conceptual in nature. Just as the spectrograms resulting from a given speaker's repetition of a particular word will differ, so too, a particular dancer's representation of the wagging run  $\omega_i \alpha_j$  may be expected to differ. Such physical variation is no key to the size of the distance or directional vocabulary.

The basis for the claim that  $\alpha$  and  $\omega$  are analog is not altogether clear. Nothing in Gould 1975b appears to substantiate the claim with respect to  $\alpha$ , although the following discussion, arising as part of the explanation for recruit accuracy in the linguistic context (see sec. 2.13), may be interpreted as an implicit objection to (75)'s division of the  $\omega$ -variable into  $m$  discrete

categories preserved under averaging.<sup>40</sup>

If recruits attended, say, 6 cycles, they would have the opportunity to average the diverging signals. If, for example, a forager whose station were 200m. away attempted to signal the distance precisely, it would have to "choose" between performing 9 waggles (indicating 190m.) or 10 waggles (214m). By performing some cycles with 9 waggles and some with 10 (4 and 2 of each, respectively, in this case), a signal accurate to within 2m of the actual distance could be created. . . . By using the scatter in the more numerous and sharply-defined sound bursts, even higher accuracies would be theoretically possible. (Gould 1975b:171)

The wagging run's  $\omega$ -variable is tied, whether or not this is its communicative aspect (see discussion, secs. 2.4.1-2), to the (integral) number of wagging movements made/sound bursts emitted; at the same time,  $\omega$  is seen to vary across SD with dance followers attending perhaps six circuits and using that information, which may well not average to an integral number of wagging movements or sound bursts, in their search. In the absence of a competence/performance distinction, it is quite proper to conclude, as Gould does, that a finite distance vocabulary is employed so as to convey via the averaging semantics an in principle infinitely fine gradation of message (SD) values. If by "graded," however, is meant a continuously varying distance-indicating capacity, then this conclusion is not warranted; averaging finitely long series of values selected from a finite alphabet can give rise to only denumerably many values. What is more, if the comprehension limit to six circuits is seen as

linguistic (i.e., a matter for the dance language grammar to deal with), then only a finite number of distance values can be so produced. Then, inasmuch as it is by no means guaranteed that  $\omega_1$  should correspond to 1 wagging movement (or sound burst),  $\omega_2$  to 2 wagging movements etc., a more finely drawn  $\omega$ -grid of finite size  $m$  ( $m$  perhaps many times the size of  $M$ , the maximal number of wagging movements, or sound bursts produced) would clearly suffice.

There are then several obvious alternatives for the lexical treatment of the distance-indicating aspect(s) of the wagging run. If, as was first suggested in sec. 3, there are finitely many conceptual (competence)  $\omega$ -categories then these categories must be preserved at the message-coding level--which means that if SD is the syntactic domain, then, unless SD is grammatically strictly a message repetition, message variation is either somehow transformationally induced from an underlying identical-message (or identical curve-supplemented wagging run) concatenation, or else between-category message codings must be marked as semantically anomalous by an averaging to a wagging run projection rule. In this view, the  $\omega$ -grid is finer grained than the integral wagging movement or sound burst variation through which the production system accesses the grammatical account.<sup>41</sup> If the maximal number  $M$  of wagging movements (or sound

bursts) produced on a wagging run and the maximal number  $p$  of circuits produced/comprehended in an SD are tied to a competence restriction, then it is clear that a finite number  $m$  of grammatical  $\omega$ -categories will suffice.<sup>42</sup>

This first alternative thus depends in a crucial way on there being least upper bounds  $M$  and  $p$ . Both assumptions are questionable.

The move from  $D$  to  $SD$ , which is to some extent sanctioned by the to-date mysterious dependency of the angular variation in a bout of dancing on wagging duration (see discussion, p. 196), ensures that if there is a value  $p$ ,  $p > 2$ . If the analogy to coordination in human language is well-taken, then the limit on the number of wagging runs per  $SD$  (whether as coordinations of message codings  $D$  or of some smaller syntactic unit), like the limit on the number of  $S$ 's (or  $NP$ 's,  $VP$ 's, etc.) which may be conjoined, is performance-based. There would, in that case, be no least upper bound  $p$  functioning grammatically.

The traditional view of wagging or sound burst duration is that it codes distance using a fatigue measure, suggesting that "[i]n theory, the dances could be performed to indicate any distance (e.g., a dance cycle indicating 100 km might last six minutes), but physiological considerations such as flight range must set practical

limits" (Gould 1975b:192). The theoretical level referred to is not unnaturally taken to be that of linguistic analysis, i.e. the grammar. If so, there is no grammatical least upper bound  $M$  on the number of wagging/sound production duration values.

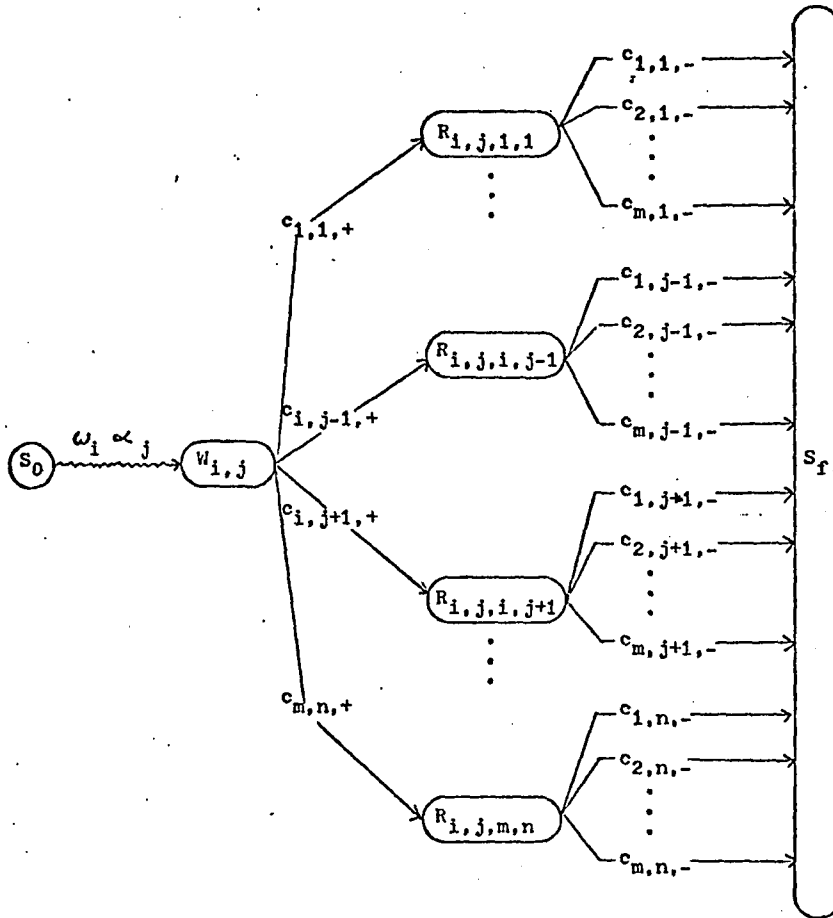
The argument against  $M$  thus depends in a crucial way on the identification of distance as what is being coded in the dance. As was pointed out in sec. 2.4.2 however, this is not necessarily the best analysis of this dance parameter. Any experimental alteration of the distance to the goal simultaneously alters the portion of the flight range to be traversed en route. This apparent "Gavagai" situation<sup>43</sup> may yield to further investigation of dance language "dialects." Should  $\omega$ -values correlate interracially in the indicated way, a flight range translation is favored; if  $\omega$ -values persist in varying interracially the issue remains undecided,<sup>44</sup> either translation requiring a dialect-dependent semantics. The distance interpretation is thus demonstrably falsifiable as an explanatory theory, though not experimentally verifiable.

The purpose of this section is not to decide this issue, which in any event awaits further experimentation, but to examine the consequences of the lexical arguments advanced. The immediate effect of the anti- $M$ , anti- $p$  argumentation is to suggest a denumerable range of  $\omega$ -values

with subscripts in the first (anti-M) case ranging across the natural numbers and in the second (anti-p) case across a denumerable set of rationals; should neither M nor p exist, the subscripts range across the full set of rational numbers, while if M but not p exists, the range is a denumerable subset of the rationals, i.e., with numerators maximally M times the denominator. There are thus three (anti-M, anti-p, and anti-M anti-p) possibilities. Their impact on the grammar is more difficult to assess.<sup>45</sup> The simplest case is anti-M, in which the  $\omega$ -grid does not stop with  $\omega_m$ , as in (75), but continues as  $\omega_1, \omega_2, \dots, \omega_m, \dots$  or, perhaps more usefully, as  $\omega_0, \omega_0\omega_1, \omega_0\omega_1\omega_1, \dots, \omega_0\underbrace{\omega_1\omega_1\dots\omega_1}_{m-1}, \dots$ <sup>46</sup> The finite state account of sec. 3 may be regarded as essentially correct (pending other criticisms, such as those of sec. 4.1) with respect to any dance type for which there is a maximal  $\omega$ -value.<sup>47</sup> Thus, only that portion of the finite state account concerned with the S-dance, namely (11), repeated here as (80), need be revised due to the anti-M decision.

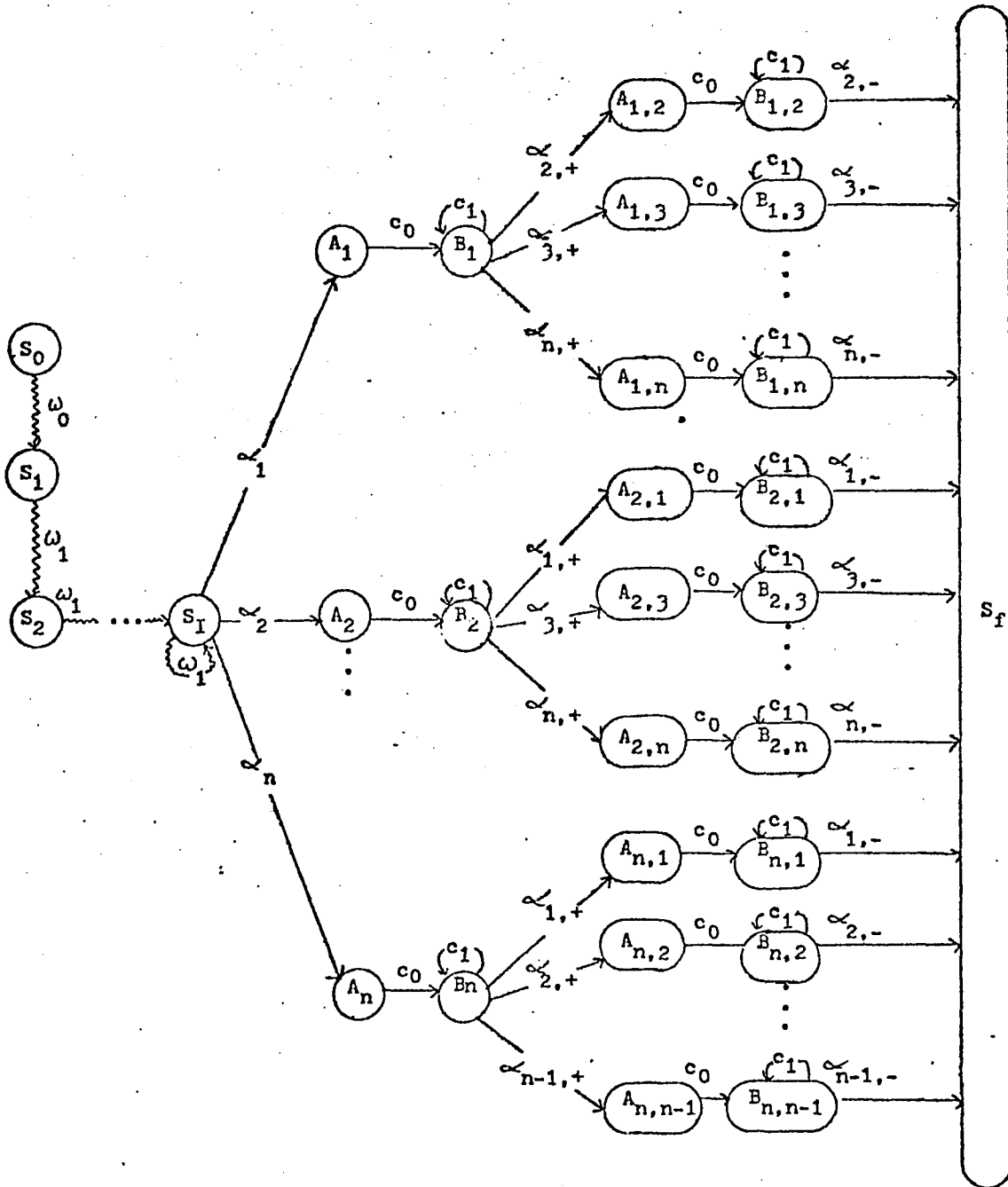
Since the only dependencies in the S-dance are between the (presumed finite--see n.45)  $\omega$ -values of the three runs, the extension of this dance type to arbitrarily large  $\omega$  distance values is easily accomplished. Each run recomposes into two features, the denumerably variable distance-correlating feature being generated using the

(80) = (11)



$\omega_0 (\omega_1)^k c_0 (c_1)^k, k \geq 1$ , notation suggested in the previous paragraph.<sup>48</sup> Then supposing  $\omega_1$  to be the smallest  $\omega$ -value suitable to the S-dance form, (80) is revised as per (81). While (80) is a schema for the S-dance type, to be supplemented by similar paths for each  $\omega_i \propto_j$  for which  $1 \leq i \leq m$  and the mirror-image additions of

(81)



sec. 3 n.3, (81) is itself the S-dance generator, to be supplemented only by its sec. 3 n.3 alter ego.

It is not obvious how to accommodate either of the two anti-p positions (anti-p and anti-M anti p) grammatically. Let  $W_0 \underbrace{W_1 \dots W_1}_{i-1}$  be the  $i^{\text{th}}$  value encountered by the diagonalizing one-one function<sup>49</sup> demonstrating the denumerability of the original set (i.e.,  $\omega$  with various fractional subscripts). The anti-p condition automatically guarantees that any sequencing of  $\omega$ -values averages to an already existing  $\omega_i$ . The problem is, to a very great extent, the reverse of that encountered in sec. 3: the averaging constraint is to some extent (i.e., with respect to the  $W$  but not  $\alpha$  parameter) met, but the syntactic dependencies are elusive, regardless of whether they are viewed as absolute (so that, for example, the wagging runs of the tailwagging dance must be identical and not merely similar) or simply statistical, with a fixed degree of variation tolerated at each stage. In sec. 3, a semantic component was suggested over and above the syntactic one. This provided a semantic anomaly grammaticality filter on syntactically generated sequences, eliminating those syntactically well-formed SD sequences which failed, on averaging, to give lexically specified  $\alpha$  and  $\omega$  values; there was an implicit objection to a cognitive "vote-tallying" process in addition to the attested averaging process. In the present case, it might

be suggested that SD sequences again be overgenerated provided only that the  $\omega$ -value is large enough for the dance type in question; then, SD's whose  $\alpha$ -variation is beyond the pale would be scrapped by a syntactic filter. In neither instance is the solution particularly attractive since the formal properties of the filter, whether semantic or syntactic, are difficult to determine, leaving the grammar difficult to assess vis à vis those which have been proposed for human language.

Finally, the cognitively simplest account is one which treats  $\omega$  as varying nondenumerably, i.e. in an analog fashion. Unlike the directional parameter, for which only a finite and nondenumerable repertoire have been motivated, the "distance" parameter is also arguably denumerable, a position Gould 1975b (see above, p. 212) apparently mistakenly characterizes as analog.

In the event that  $\alpha$  and  $\omega$  (or both) are nondenumerable, so is the dance language. Then, following the argument sketched in n.39, the dance language is outside the Chomsky hierarchy of (36), syntactically different in kind rather than degree from human language.

It is this claim which is the focus of the next section.

#### 4.3 Syntactically Different in Kind Rather than Degree

The investigative approach which has been developed

here is in marked contrast to, for example, that of Hockett and Altmann 1968, which contents itself with providing a list of parameters or design features to aid in the description of communication systems. More than half of the design features specified are extra-grammatical. These include an insistence on the employment of a vocal-auditory channel, broadcast transmission and directional reception, and rapid fading of the signal. While the status of some of the other criteria is unclear, these are surely not linguistically significant, since filtering the set of English sentences through a simple (e.g., letter for letter substitution) written code robs it of all three properties. Others of these features, while certainly true of human language, appear to be of minimal cognitive import. The requirement that all transmitters also function as receivers and vice versa (interchangeability), that the signal not have any purpose other than that of communication (specialization), that the signal be at least some of the time spatiotemporally removed from the stimuli provoking it (displacement), are design features of this type; in each case, the honey bee's dance language appears to meet the criterion (see, for example, Thorpe 1974:70<sup>50</sup>).

The displacement design feature is apparently intended to have cognitive significance, but the work of Dethier on the "dance" (fidget?) behavior of flies (see

p. 198) shows just how poor a cognitive indication this in fact is. Dethier 1957:334 reports that flying is an inhibiting factor for this species, increasing the maximum delay between the sugar stimulus and the "dance" response from 1 minute to 5. While the forced flying induces spatiotemporal displacement, it is not thereby evidence against a simple stimulus-response model. In the case of the honey bee, the inevitable spatiotemporal delay between the visit to the goal and the dance in the hive is therefore not particularly impressive with respect to a cognitive analysis since, again, that delay is mediated by flight activity. What evidence there is for displacement as a cognitive parameter met by the dance language is provided by the dancer who is, so to speak, reactivated by a light source introduced into the dark hive, provided it can be shown that since her last bout of dancing, a reactivated dancer may have been engaged in the full spectrum of activities ordinarily pursued in a forager's off-hours (see Lindauer 1967:20-21). Even after it has been so clarified, the design feature is of limited cognitive significance.

A similar situation obtains with respect to the design feature of Complete Feedback, which requires that the transmitter have as much access to his output as the receiver does, presumably so as to help ensure that the transmission is being monitored and so is not merely an

involuntary "knee-jerk" reaction. Interestingly, in the dance language case, it may be easier to determine whether there is self-monitoring than whether there is complete feedback. This is so because of the difficulty encountered in determining just what the dance attender is in fact sensorily attending to: there are, for example, numerous distance correlations in the dance (the duration of wagging/sound production, the  $\alpha$  variation in an SD, the number of wagging movements or sound bursts emitted, etc.) with no real evidence<sup>51</sup> to date as to which is/are the communicative aspect(s). On the other hand, given the traditional view that each wagging run is an attempt at the specification of the site of the goal, the fact that SD variation about the goal's coordinates cancels out, leaving the average of the danced values roughly on target, suggests error-correction and thus indicates self-monitoring.

The most mysterious of the criteria is Learnability, defined simply as: "A speaker of a language can learn another language" (Hockett and Altmann 1968:64). Why should a property of a system (i.e., its linguistic status) depend on a property of the user of that system? Would English be any less a language if man's language acquisition device were such that English was the only system it was possible to acquire?<sup>52</sup>

The design feature Prevarication requires that the organism be able to "say things that are false or

meaningless" (Hockett and Altmann 1968:84). Not simply that for a given context there exist false or meaningless SD's, but that these in fact occur in that context. This is notoriously difficult to demonstrate, the initial reaction to apparent examples of this kind being just that, that they are only apparent cases of prevarication, in actuality constituting evidence that the system is not yet fully understood or, perhaps, evidence that the transmitter is suffering from a physiological disturbance of some sort.

This criterion is moreover vulnerable to the same criticism as was Learnability, viz., that a property of the system's user is being confused with properties of the system itself. Consider the case of a Truth-Teller, who (perhaps even due to a genetic block) eschews both falsehoods and semantically anomalous sentences. If Hockett and Altmann are correct in demanding Prevarication of language, then in identifying that individual as a Truth-Teller, presumably on the basis of his speech, we are discovering we do not share a language. This is absurd.

Bennett 1964 proposes many of the same linguistic criteria as Hockett and Altmann 1968<sup>53</sup> but strongly emphasizes a parameter which can be viewed as an extension and elaboration of specialization, the requirement of grammatical consciousness. The intention is to discriminate between regular<sup>54</sup> and rule-guided behavior, justifying

"the move from descriptions which are rules to descriptions which refer to rules" (Bennett 1964:21), ultimately deciding whether "the bees are in the grip of the rules ... [or] the rules are in the grip of the bees" (Bennett 1964:17). The criterion: "if the dances are symbolic then of most dances it must be true that the bee performs the dance because it is in the presence of other bees which will be informed by the dance as to the whereabouts of food" (Bennett 1964:45). Specialization requires that the behavior have no purpose (selectional advantage, in evolutionary terms) other than the communicative one; grammatical consciousness requires that the individual communicator intend to communicate and have no other intention in executing the behavior in question. It is claimed that truly linguistic behavior subsumes intelligence but not rationality, this last being the topic of Bennett's essay. The bees must be "intelligent about the circumstances under which they refrained from dancing, or relayed one another's dances, or refrained from acting on the dances of certain bees, etc." (Bennett 1964:48), the epithet of "intelligent" being reserved for responses which do not have a sensory-causal explanation.

The following (Prevarication-influenced) scenario is developed in Bennett 1964 as an illustration of what would justify the imputation of intelligence. It is argued that the required awareness of the rules as rules

is demonstrated by (for example) a dance attender's executing a particular denial maneuver when faced by a dance featuring a breach of the rules; the breach which is of interest is that which arises when a dance attender arrives at an advertised site which (from the bee's sensory point of view) utterly lacks food-traces, i.e. utterly lacks any (bee-sensible) indication that there was recently something of value at the site. Then, provided the dance attender can be conditioned to widen her notion of food-traces to embrace various non-biological experimenter-correlated factors, she is intelligent and the dance language is a language. Bennett (1964:48-49) concludes by remarking that his (Prevarication-influenced) scenario shows that the bee is intelligent about what to dance as well as whether to dance. The implication is that while evidence of the sort sketched in the previous paragraph is evidence that the dance language is rule-guided, evidence of message consciousness (and not merely grammatical consciousness) is a prerequisite for a semantic rule.

In spirit, the criterion is perhaps more like the grammatical design feature of Reflexiveness. "In a language we can communicate about the very system in which we are communicating" (Hockett and Altmann 1968:64) than any of the others, in that it decides the linguistic status of the system on the basis of evidence that the system

does or does not allow for its utterances to be linguistically challenged or approved. This in turn makes the status of the system dependent on its expressive power vis à vis the construction of a linguistic theory; if the dance attender cannot either castigate the dancer's behavior as ungrammatical, inappropriate, ill-formed, etc., or approve a dance as grammatical, appropriate, well-formed, etc., the dance language is not a language.

Without such an overt linguistic criticism to point to, the experimenter must for simplicity reasons regard the silent dances Esch 1961 investigated, which do not arouse recruits, not as rule-breaching examples, but as indications of a physiological disturbance of some sort. The fact that such dances are ignored is presumably to be tied either to the dance attenders' impression of the dancer as deviant, as opposed to a recognition that the dance is deviant, or, perhaps, to the dance attenders' sudden recognition that the bee isn't dancing after all. Either "explanation" will do.

The arguments that ensue between real-estate scouts who advertise different sites to the swarm cluster (see Lindauer 1967:34-54, von Frisch 1967a:269-73) are also easily dismissed. Each scout (so far as is known) reports the results of her own sensory comparison of the sites; the dance can only evoke a search response, not a danced reply. Consequently, the danced agreement which

must be reached before the swarm cluster decamps is not linguistic but sensory agreement. And if the dance should prove to be relayed in some circumstances, as when there is an enormous difference in the vivacity with which different sites are advertised, then surely that is also evidence that the bees are less than intelligent about their communication system--for, lo and behold, the dances are "catching."

The thrust of Bennett 1964 is that unless the organism can be shown to have an embryonic linguistic theory, it cannot be argued to have a language and unless it can be shown to have an embryonic science, it cannot be shown rational. The scenario for rationality is an extension of the above-described Prevarication-influenced scenario for linguistic status. It is argued that the indication of rationality is the ability to linguistically code statements of the form "x because of evidence e" and their denials (taken, as usual, to be roughly of the form "e at least once without x"), with the speech community enjoying a fair degree of latitude in their response to an argument (see Bennett 1964:69-71).

Thus, without a bee-scientist, there can be no question of rationality--without a bee-linguist, no question of linguistic competence. Short of a bee linguist, the only pro-language evidence that can be discovered is evidence that dancers (intentionally) lie and can be

accused of lying. In a society for which lying can have no possible motivation, a lying bee would be behaving irrationally (Gould 1975b:204). Granted that the only topic the dance language is known to treat is distance and direction from the hive to a goal (and, perhaps, the value of the goal given current hive-conditions, which is coded by the dancer's vivacity and is here assumed to be paralinguistic), must it follow that the system used to code information on that topic is not a language? A criterion such as this one denies that the honey bee is a language-user on the grounds that its biologically determined interests differ from man's.

The Bennett criterion of grammatical consciousness is left open to a number of interpretations varying in strength. The weakest of these, that the system be employed in such fashion that intelligence may reasonably be inferred from the pattern of "language" use, is arguably met by the honey bee; for, she dances only in the presence of other bees (von Frisch 1967a:36), and with experience learns to extrapolate, and linguistically compensate for, solar azimuthal change. If this is, in fact, intelligence as Bennett uses the term, the dance language is thereby apparently to be accorded linguistic status as a rule-governed communicative system, for which grammatical investigation is warranted.

If, however, each hypothesized rule or rule-type (as syntactic vs. semantic, for example) must be

justified by a similar demonstration of the dancer/observer's awareness of rules as rules, as is certainly suggested by Bennett's choice of scenario, the dance language is of dubious grammatical complexity. What Bennett fails to note is that grammatical consciousness, given this strong interpretation, also renders human language only a very small degree of grammatical complexity, all the while leaving the central issue of distinguishing the well-formed sequences of (say) English sounds, morphemes, etc. unanswered, and essentially unanswerable. What's more, the fact that the child, the language acquirer par excellence, displays little if any (strong or weak) grammatical consciousness during the critical period, is demonstrably insensitive to the rule-presentation variety of formal instruction for a second language, would force a Bennett-type analysis to the untenable position that the grammar springs virtually fully-formed into the mind of the mature speaker/hearer, for it is only with maturity, with grammatical consciousness, that grammatical rules may be posited for the language-user.

This is not to suggest that English is regular rather than rule-guided--to be sure, the question only arises for heritable systems with a cognitive appearance, such as the dance language. The problem is that Bennett's proposed criterion, grammatical consciousness, can not be used to decide the regular vs. rule-guided question for

such heritable systems unless, at least, it gives the correct (rule-guided) response for such clear cases as that of acquired systems such as human language. Grammatical consciousness does not meet this level of minimal adequacy.

Kroeber 1952:753 addresses the issue in sign vs. symbol terms, quite simply asserting that a heritable system uses signs, not symbols; symbols must be assigned their specification by convention and not by genetic coding. This then conflates the regular vs. rule-guided problem (to which we return on p.240, below) with the design feature of Cultural Transmission or Tradition. Not even the acquired incrementing with solar motion aspect of the dance repertoire presents a clear case of Cultural Transmission, for that would require that an individual dancer's learning to increment in an SD be contingent not only on her experience of solar motion but also the observation of at least one case of SD incrementing with solar motion by a linguistically mature dancer/observer.

Cultural Transmission, common to most list of criteria, is typically motivated in part by the mistaken notion that no very substantial demand can be made of a system without an acquisition parameter. Mistler-Lachman and Lachman 1974:871 reveals a particularly heavy bias of this sort: "by the weak criterion, man, computers, and some chimpanzees use language . . . weak equivalence requires

only . . . labeling, syntax, and semantics." Their strong criterion is strong indeed, requiring that the language user not only demonstrate all universal linguistic patterns, but have acquired his language, in acquiring it passing through the same developmental stages as the child (making similar errors), by means of similar neurological constructs.

The claim that language is an exclusively human achievement has empirical content only insofar as "human" is not built into its definition. Mistler-Lachman and Lachman's strong criterion defines a notational variant of a human language-user and then, arbitrarily and without argumentation, equates it with a human-language user. The prevalence of the practice (see Griffin 1976) does not make it scientifically palatable.

The Cultural Transmission design feature also seems to be motivated by the central importance of distinguishing the innate from the experiential component of the human language acquisition process. This motivation, however, is more apparent than real. Consider, for example, a species for which a particular human language (say, English) were genetically coded.<sup>55</sup> Hockett and Altmann's Cultural Transmission design feature will mark this as a non-language, leaving man the lone language user; a non-Cultural Transmission approach, such as that developed here, would accord this system full linguistic status, characterizing as a (boringly simple) notational variant of a human

language--by which is meant a system that the normal child acquires as per Mistler-Lachman and Lachman's strong criterion, above--leaving man not the lone language user, but the lone language acquirer. Cultural Transmission plays a role in that linguistic universals are isolated as that part of the system for which tradition is not an explanation; these in turn define the dimensions of the possible input to the child's Language Acquisition Device. It is unclear what if anything is gained by reintroducing the Cultural Transmission parameter directly into the characterization of that input as language.


The rationalist transformational revolution in linguistics has focussed attention on the mathematical complexity of the grammar, viewed now as a cognitive window on the species. Despite the fairly extensive lists of linguistic desiderata which have been proposed as the yardsticks against which animal communication systems are to be measured, this issue remains as the essential, and only theoretically principled one with regard to the linguistic status of the honey bee's dance language. Mistler-Lachman and Lachman's distinction between product and process criteria is a distinction between linguistically principled and unprincipled measures; any system requiring labelling and grammatical complexity of the kind found in human language, which are the linguistic universals, must be designated language.

Hockett and Altmann 1968 propose six grammatical design features; the dance language is generally considered to meet the criteria of Openness and perhaps Semanticity (although this has been questioned), but not Arbitrariness, Duality of Patterning, Reflexiveness (see discussion, p. 226), or Discreteness.

The dance language is open, productive, creative if the repertoire is not memorizable (if "[n]ew messages are coined freely and easily, and, in context, are usually understood"--Hockett and Altmann 1968:63), i.e. if the repertoire is infinite; as previously discussed, this follows from an infinite lexicon and/or unboundedly long SD's as the syntactic domain. This, as Limber 1977:282 points out, is in itself a very minimal demand: "A language whose nonfinite nature was due only to something like recursive intensification, as in The banana was very, very, ... very good, would be of relatively little significance." As will be seen, none of Hockett and Altmann's design features guarantee the right sort of infinity; what Limber 1977:282 defines as "the projective aspect of human language ability," described in the following terms: "[u]sers of any human language are able to syntactically project novel yet appropriate linguistic expressions onto any of an almost unlimited number and variety of concepts."

Limber 1977 is particularly cagey as to what aspects (syntactic, semantic, or both) of the infinity make

it "the right sort," although two syntactic universals are explicitly mentioned: constituent structure and sentence embedding.<sup>57</sup> A syntactic approach will note that recursive intensification yields to a finite state treatment

(with a loop ) , while a semantic approach will have to specify the crucial message coding properties of human language.

Katz 1976 exemplifies the semantic approach, first proposing three essentially non-semantic criteria of unboundedness (openness), stimulus-freedom, and appropriateness, only to observe that "it does not seem likely that any or all of them constitute the uniqueness of natural language" (Katz 1976:35). Stimulus-freedom and appropriateness are dismissed as performance-oriented<sup>58</sup> criteria, appropriateness moreover on the grounds that schizophrenic dissociation is not a linguistic deficit. Syntactic unboundedness is, as already noted, too modest a demand. The semantic universal which is to characterize human language is the effability thesis: "Every proposition is the sense of some sentence in each natural language" (Katz 1976:37). The criterion entails full, mutual translation between natural languages: "For any pair of natural languages and for any sentence S in one and sense  $\sigma$  of S, there is at least one sentence S' in the other language such that  $\sigma$  is a sense of S'" (Katz 1976:39).

Effability may be thought of as the semantic counterpart to syntactic unboundedness. The syntactic criterion hypothesizes that the set of syntactically generated objects will not vary in cardinality across natural languages, while the semantic criterion hypothesizes that the set of semantic representations is not only equi-numerous but identical across languages; both are claims about the nature of the output of grammatical rules. Neither claim, and this is crucial, effectively describes the grammatical constructs responsible for ensuring the appropriate output property.

Unlike unboundedness, however, which is motivated on grounds of syntactic rule generality, effability may reasonably be expected to differentiate human from animal language, accounting for man's--but not necessarily Language's--uniqueness (see discussion, p. 229).

It seems that the crucial linguistic question is whether communication systems (or, more generally, the cognitively based systems of different species) are structurally equivalent. It is to be expected that structural equivalence will include both syntactic and semantic parameters,<sup>59</sup> but the focus of these must be on rule types and grammatical organization, rather than on rule output.

Hockett and Altmann 1968 supplement Openness with design features which, again, are primarily output constraints, and hence quite suspect as a guarantee of

linguistic status.

Arbitrariness requires that there be no "physical or geometrical" resemblance between the message and its coding; Hockett 1960:428 and Thorpe 1974:70 accordingly judge the dance language to be iconic, while Kroeber 1952:755 finds the negative geotaxis/positive phototaxis alternation to be an apparently arbitrary, symbolic element which is (from his point of view, disturbingly) genetically coded. A species-independent, tradition-independent notion of arbitrariness may well prove impossible to specify, but it seems safe to describe the dance language, on balance, as iconic, unlike human vocal language. What is not-clear is the relevancy of this observation to the linguistic evaluation of the system. Why should this be viewed as criterial?

Surprisingly, Duality of Patterning, the only grammatical design feature which is not an output constraint, may in fact be met, courtesy of the transitional dance form. That is,  $\omega_i \alpha_j c_{i,j,+}$  may directly yield the goal's coordinates as, say,  $\omega_i \alpha_j \oplus 3$  by the following comprehension process:  $i$  is a (for example) six-degrees-of- $\alpha$ - deviation distance value which, since the consequent return is counterclockwise, implies that the observed  $\alpha_j$  is then  $6/2 = 3$  units off to the left, giving the goal's coordinates as  $\omega_i \alpha_j \oplus 3$ . According to this comprehension model,  $\omega_i \alpha_j$  and  $c_{i,j,+}$  are themselves directionally meaningless<sup>60</sup> (as  $\omega_I \alpha_j$  is not, if  $\omega_I$  denotes a tail-

wagging or S-dance distance value), meaningful only in combination. This is admittedly not the traditional view, and would additionally require evidence both that  $c_{i,j,+}$  is a linguistic category (curved runs were ignored in sec. 4.1.2) and that the distance value is read off  $\omega_{i \propto j}$  and not  $c_{i,j,+}$ , which would otherwise be the sole meaningful dance element for this dance type--an unfortunate result since it certainly plays no semantic role with respect to the S-dance or round dance and need not play a role with respect to the tailwagging dance. The intention here is not to claim that Duality of Patterning holds but, more modestly, that it does not obviously fail to hold. This leaves Semanticity and Discreteness, which together force a return to the regular vs. rule-guided question.

Semanticity is the demand that "some linguistic forms have denotations" (Hockett and Altmann 1968:63). Kroeber 1952:753, in drawing a tentative acquisition-based distinction between sign and symbol, grounds that distinction in one between subjective and objective communication, the (objective) symbol being able to "convey information on other matters than the condition of the communicating organism." The connection is clear: if regular, non-symbolic, the dance language codes fatigue, while if rule-guided, symbolic it codes distance or flight range percentage, gauged by a fatigue measure, as outlined in sec. 2.

In either case, it is not simple fatigue, but fatigue compounded by a perceptual filter weighting the outward flight. Semanticity is then satisfied only by the objective, rule-guided interpretation of the system. Again, this is not a point which can be settled by experimentation. The very considerable evidence of (perceptual?) integration of cross-model stimuli in determining danced values is suggestive, but need not be read as evidence of multiple sensory estimates of the objective truth to be conveyed.

Bennett's grammatical consciousness parameter patently fails to discriminate between these alternatives. The single clearest case of grammatical consciousness is, ironically enough, that of the "linguistically" trained chimpanzee. It is a curious property of this awareness that, in clearly rule-guided (i.e. acquisition) cases, it is most easily demonstrated for learning in a domain for which the subject is biologically disadvantaged. Where there is an interesting cognitive substrate biologically guiding the learning as, in the case of the child, there are innate linguistic universals, there is little rule awareness.

If the dance language is regular, the dancer is simply a physical transducer, helplessly reporting her internal state, which correlates with observable, external conditions in an unusually obvious way, thereby "looking

symbolic." There would, in that case, be no cognitive weight to the system; in the terms of secs. 2-3, no competence/performance distinction, because no competence.

This observation, far from undermining the present analytical investigation, provides an approach to settling this instance of the regular vs. rule-guided debacle. Quite simply, should the positing of a competence idealization (any competence idealization) contribute to an understanding of the communication system, it is thereby justified, as is the interpretation of the dance language as rule-guided.

It is an empirical issue, manifestly not to be decided by a simplicity, no-cognitive-structure-here criterion which in the present context is no more than thinly veiled chauvinistic contempt for infrahumans, a matter of intellectual fashion rather than argumentation, to quote Bennett.

The digital approach of sec. 3 provided a notation in which to outline the more salient structural dependencies of the dance behavior, incidentally drawing the reader's attention to features which are difficult to explain away as by-products of a physical transducer at work (as, for example, the curver run's circumscribed variability). Granted that the dance language's status as a rule-guided system depends on specifying an observationally, descriptively, "explanatorily" adequate grammatical account, it

must be admitted that thus far, the attempt has not been an unqualified success. No digital account explored has succeeded in making sense of the system as a whole, explaining (for example) the 60° limit to  $\alpha$  variation, the parameters of dialect variation, or even the apparently pointless variety of dance types, much less their confinement to a distance range.

The final design feature is that of Discreteness, the main concern of sec. 4.2. The hypothesis that the dance language is continuous, syntactically different in kind rather than degree from (discrete) human language dates, at least, to Haldane and Spurway 1954:263. In recent years, it has surfaced in Hockett 1960:428; Thorpe 1974:70; Janda 1977:147-48; and Chomsky 1975b:301, generally without any supporting evidence or argumentation, as if it were derived by simple observation of the dance behavior. This is absurd, all behavior being continuous at some level of analysis, unless of course the system is regular, in which case there is only one analytical level. This appears to be the force of Chomsky's (1975b:301) remark that "an animal<sup>[61]</sup> can operate on the principle of the speedometer, producing a potentially infinite, in fact in principle continuous, set of signals as output in response to a continuous range of stimuli-- . . . continuous, in the only sense in which any physical signaling system can be regarded as continuous." It may well also

be the basis for Hockett 1960's and Thorpe 1974's decisions in this regard.

Janda 1977, on the other hand, makes the more serious methodological error of deciding on a continuous lexical representation (for both the  $\alpha$  and  $\omega$  variables, in fact) essentially on the grounds of cognitive simplicity.<sup>62</sup> Again, the issue is an empirical one, tied to the levels of adequacy demanded of a grammatical account.

That error is rectified here, in sec. 4.4, with the presentation of an analog analysis accounting for the structural features which were stumbling blocks for the digital approach. The discussion opens with an outline of the grammatical alternatives thus far pursued, and the aspects of the behavior which remain mysterious; sec. 5 outlines the account's experimental implications.

#### 4.4 The Search Area Hypothesis

The popular view of the dance language as a regular, stimulus-response system, iconically coding the dancer's flight to the goal, fails to so much as hint at the rich structural substrate of the behavior. It blindly ignores as physiological interference (of unknown origin) the fact that the dance is not a simple buzzing run--or, what would be even more convincing, a series of runs recapitulating that dancer's successful search history.

This view is a view about the stimulus-response

nature of the message set, not the dance language, and as such is entirely innocuous. Its extension to the dance language is, as sec. 4.3 indicated, far from established.

The dance language presents a series of patterns for which there is only very minor genetic variation. This variation is concerned solely with the form of the transitional dance and (of a distance interpretation is insisted on) the precise figures assigned as the range of distance values appropriate to a given dance type. Every dancing race identically orders the dance forms with increasing distance. Every dancing race has the round dance, a transitional form gradually approximating the tailwagging form, and the S-dance. Why should there be different dance forms? And what ties the dance form to its position in this universal near-the-hive hierarchy? Why is SD angular deviation limited to  $60^\circ$ ? Why does SD angular deviation correlate inversely with the  $\omega$ -value? Why are the round and S-dances so much "sloppier" than the others? Why should there be dances which indicate direction without distance?

The previous sections have experimented with a number of competence/performance titrations without, however, managing to provide real insight into any of these issues. The assumption in sec. 3 of a finite  $\alpha$  and  $\omega$ -range allowed for a finite state treatment which, complete with categorizing symbols  $I_W$ ,  $I_D$ , etc. (see sec. 4.1.2.), is

arguably descriptively adequate and amenable to at least some versions of the SD claim. Even if the goal of descriptive adequacy were to mandate (as described in sec. 4.1.2) spread transformations or semantic averaging rules, the resulting grammar would not be improved with respect to the issue of "explanatory" adequacy.

If the  $\omega$ -range were denumerable, but  $\infty$  finite, the treatment would depend to some extent on the source of  $\omega$ 's denumerability. If the denumerability were to be produced, in part or in whole, by the averaging process (whether or not averaging is itself grammatically coded), then it is easily accommodated by a simplification of the finite state base component, so that integral averaging output is no longer ensured. If the denumerability is due to the in principle unboundedly long wagging runs which, for example, Gould 1975b:192 (see p. 214) argues for,<sup>63</sup> then the solution is in terms of an extension of the S-dance to a denumerable series of wagging runs as per (81). The only problem that remains is in connection with a denumerable set of  $\omega$  values arising from unboundedly long wagging runs, length measured in an integral number of wagging movements/sound bursts, as:  $\omega_0, \omega_1, \omega_2, \dots$ , taken in conjunction with the SD claim and an average-to-an-integer grammatical constraint. The obvious approach: a filtering device. Again, only descriptive adequacy (at best) is achieved.

The more modest cognitive claim that  $\omega$  or  $\omega$ , or both, are nondenumerably infinite thus merits thorough investigation. With the shift to a nondenumerable range, the grammar for the dance language moves outside the Chomsky hierarchy, and is syntactically different in kind rather than degree from human language. The syntax can not be a scheme for generating the sentences of the language, but must instead define the properties of those sentences so as to distinguish, without generating, the grammatical from the ungrammatical sequences. Given the level of descriptive achievement of the various digital accounts explored, it is the level of "explanatory" adequacy which must motivate the analog approach. It is moreover clear that a move regarding the finite state bracketing base (with or without spread rules or semantic component) as essentially correct, but interpreting the subscripts referentially rather than indexically<sup>64</sup> to allow for the nondenumerability of the vocabulary is illegitimate; such a position would be tantamount to admitting that there was no linguistic advantage in either hypothesis as to vocabulary size. This urges the nondenumerable hypothesis as the simplest one in the circumstances, a move which is then often misinterpreted as somehow establishing the linguistic status of the dance language as cognitively inferior, when in fact it begs the question. What is needed is an analog "grammar" which will answer some of the questions raised in

the opening paragraphs of this section.

The key lies in drawing the competence/performance distinction somewhat differently than it has been to date. The unifying property of all the accounts thus far considered is their identification of the message of an SD with the point at which the dancer had been feeding (drinking, inspecting as a nesting site, etc.). This blind acceptance of the linguistic reality of the stimulus input has had the effect of tying the various digital accounts very tightly to a stimulus-response model, whose "cognitive contribution" of, for example, finite  $\alpha$  and  $\omega$  ranges might well be read as sensory threshold phenomena. The single serious disagreement with that model has been in terms of the transitional dance (see discussion, pp.188-91), which has been viewed throughout as something other than an error-ridden tailwagging dance form. That is, the "error" is taken to be in need of linguistic explanation. It is then no accident that the transitional dance became the focus of so many of the power-of-the-grammar arguments launched within the discussion of the digital account. Gould 1975a's patch theory (see discussion, pp. 197-98) is an attempt to explain the patterning of such "errors" over SD's of all dance types--it is to be tied to the profitability radius of the food each race favored, or the difference in search strategies or colony spacings each race employed. This evolutionary emphasis is of course a

direct consequence of the error label assigned the phenomena; the notion of historically different habitat/colony/search-pattern configurations for each race is a consequence of the strict distance interpretation of  $\omega$ , which then creates the problem of explaining why the patches are racially variable in size.

Attacking the problem within the larger context of a competence/performance distinction suggests an alternative explanation, couched in cognitive rather than physical terms. It will be shown that this notion of the message encoded by an SD as a patch of fixed radius about the goal, which will be referred to as the search area hypothesis, answers each of the questions raised at the beginning of this section. The bee's cognitive contribution, which is to justify the imputation of rules, is this extension from the point to the patch, whose genetically fixed radius is a biological by-product of the physiologically determined flight range of the species/race/stock of honey bee. The competence/performance distinction proposed is a natural outgrowth of the emphasis on linguistic form which dictated the earlier treatment of the transitional dance as something other than an error-ridden tailwagging dance.

Gould 1975a's patch theory is an attempt to find an adaptive value for the observed "error" in the message coding, where the message is understood to be the distance

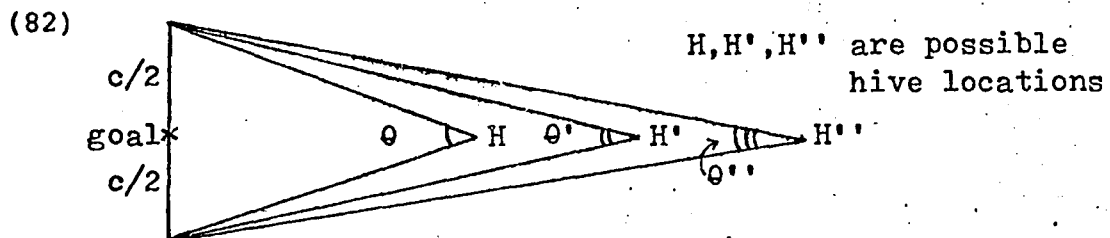
and direction from the hive (or dance site) to the location previously visited by the dancer, properly idealized to a point. SD has the selective advantage of casting recruits about the vicinity of the (point) site exploited by the dancer. Given this view, Gould's (1975a:692) remark that "[t]he area around the hive for which round dances are performed might be thought of as the first such 'patch'" is distinctly odd.

A comparison between the round dance range of use and the size of the area which each SD attracts recruits for, is in no way motivated by the theory in question. While the size and general contour of the recruit distribution after a round dance is accounted for, the fact that the same dance behavior is provoked by goals in all directions of the hive is as mysterious as is the centering of this "first patch" about the hive. Far from being explained, the round dance is a theoretical anomaly.

The search area hypothesis, on the other hand, in identifying the message of the SD with an area to be searched, the task of the dancer with the linguistic coding of the patch, does in fact motivate a comparison between the round dance range of use and the search area. For in that case, it is to be expected that the cognitive extension from the point source to the patch, when done at a point close to the hive, would lead to a performance problem; should the patch include the hive itself (the very food

supply which foraging is designed to maintain), it is to be expected that locative coding would collapse, allowing the dancer to report only that something had been found. The round dance is not a type of dance at all. It has no semantic coding and by virtue of that fact, conveys the message by default, as it were.

Assuming, with Gould (1975a:692) that "the reduced angular scatter compensates for increasing distance," with  $c$  the constant distance error induced by the angular scatter (see (82)), immediately suggests  $c/2$  as the range



of use of the round dance. That is, if the point-to-patch cognitive extension produces a patch of diameter  $c$ , then the genetically imposed radius of extension is  $c/2$  and performance problems should arise at feeding sites where an attempted  $c/2$  point-to-patch extension encompasses the hive, i.e., for feeding sites at most a distance of  $c/2$  from the hive

It has been claimed that the search area hypothesis has the advantage of explaining the round dance's cross-racial persistence for a fixed range about the hive. The nature of that explanation requires a particular formulation

of the search area hypothesis, in which the message coded by an SD is the set of points lying anywhere within a circle of rationally fixed radius  $r$ , centered on a goal. The description of the boundary as a circle follows from the lack of directional dependence of the range of use of the round dance; the inclusion of all interior points follows from the use of the round dance for a range of distance values rather than for goals at a fixed distance. There can, of course, be no round "dance" distributional evidence relative to the question of whether or not the search area boundary points are included in the search area, although it seems bizarre to allow points arbitrarily close to the boundary without allowing the boundary itself to be part of the search area. Finally, the goal is at the center of the search area even as the hive is at the center of the "provokes a round dance" area. It then follows that search areas constructed about goals less than or equal to  $c/2$  units from the hive will include the hive and so be problematic, causing the round "dance" behavior.

The set of semantic representations for the dance language can then be defined as a set of solutions of equations--the equations characterizing the interiors plus boundaries of circles of fixed radius. It seems likely that this fixed radius, given in flight range percentage rather than pure distance terms, will prove to be rationally invariant.

The coordinate system most congenial to the dance language is that of polar coordinates,<sup>65</sup> with the hive at the pole or origin (0,0). If  $(r_0, \theta_0)$  is the center of the  $c/2$  - radius search area, i.e., if  $(r_0, \theta_0)$  is the goal, then the search area (interior plus boundary points) is given in (83). This is the canonical form for the

$$(83) \quad \left\{ (r, \theta) / \left[ (r \cos \theta - r_0 \cos \theta_0)^2 + (r \sin \theta - r_0 \sin \theta_0)^2 \right] \leq (c/2)^2 \right\}$$

semantic representation of an SD. The set of semantic representations is given in (84). The R-condition holds

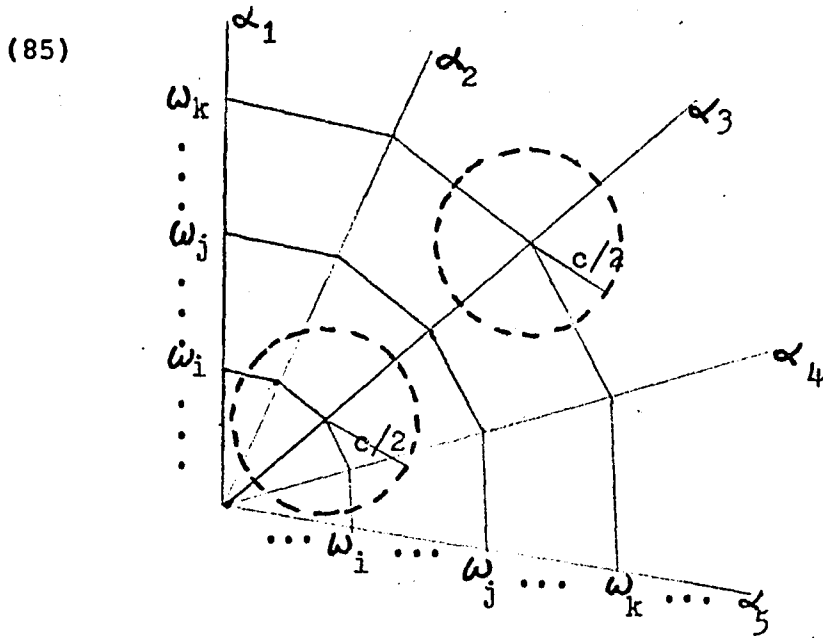
$$(84) \quad \left\{ \left\{ (r, \theta) / \left[ (r \cos \theta - r_0 \cos \theta_0)^2 + (r \sin \theta - r_0 \sin \theta_0)^2 \right] \leq (c/2)^2 \right\} / r_0 < R \right\}$$

only if a percentage of flight range framework allows for a racially independent search area radius of  $c/2$ ; this is strongly suggested by a cross-racial correlation between dance type and wagging duration. In that case,  $R$  (which would correlate with the same extreme value for  $\omega$  across races) may be read as 100% of the flight range and the other  $r$ -values translated accordingly.

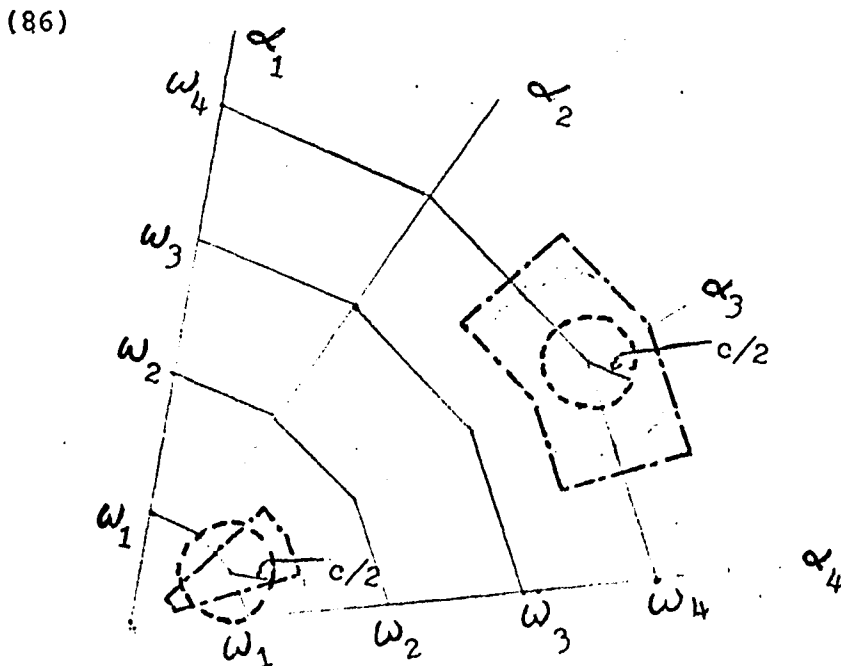
The syntactic coding must determine which of the nondenumerably many possible messages in (84) is being sent. Now, that message could conceivably be coded simply by indicating the center  $(r_0, \theta_0)$  of the search area, i.e. simply by referring to the location of the goal. The kinds

of errors that are made in so referring to the goal are errors having to do with confusing, in performance, the lexical representation of the center of the search area with the lexical representations of other search area points; this explains (à la Fodor 1966) the inverse relationship between the degree of SD angular deviation and the distance of the goal--as a performance effect at the semantic level. In that case, it might even be argued that the lexical vocabulary is countable, and correlates with a countable subset of (84) specifying search areas about the lexically specifiable goal locations, which countable subset is in fact the set of semantic representations.

In other words, it might be argued that the search area hypothesis, construed as a theory about the semantic interpretation of SD's, is independent of the centrally important syntactically-different-in-kind-rather-than-degree issue. This is to miss the point. For, consider the four alternatives (see sec. 4.2) to a nondenumerable vocabulary; each one requires that there be only finitely many  $\alpha$ -values,  $\alpha_1, \alpha_2, \dots, \alpha_n$ . Even if denumerably many  $\omega$ -cuts mediate those indicated in (85), syntactic coding will give better search area approximations near the hive, where the  $\alpha$  spokes are closer together, i.e., where that search area could have been centered on a larger number of lexically specifiable  $(r, \theta)$  pairs. Indeed, if  $\omega$  is allowed to get sufficiently large relative to the spacing




of the  $\alpha$  values, then the degree of angular uncertainty as to goal location is greater than the "uncertainty" contributed by the search area semantics (see (86)).




The point-to-patch cognitive translation would make a contribution whose recruiting effect could not be generalized across goals.

More importantly, regardless of the cardinality of the vocabulary employed, a syntax coding the center of the patch fails to provide any basis for explaining the full variety of dance types. The round dance is accounted for, as is the gradual "closing" of the transitional dance to the tailwagging shape. But there is no reason to expect anything but a sharp shift between the round "dance" and wagging transitional forms (no reason to expect the

gradual , etc. approximations to

the  dance, which is in any case an error for

, no reason to expect S-dances for such a large range of values near the foraging limit.<sup>66</sup>

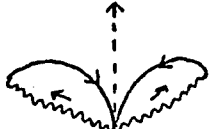
Of the vast number of syntactic strategies that imaginably code the search area by other means than a reference to the goal, only one directly reflects the evidence originally offered for the search area hypothesis, viz. the patterning of angular deviation with distance. That strategy views directional variation, but not distance variation, as an integral part of the syntactic structure of an SD. Now, the semantic representation an

SD codes can be recovered either by knowing that a particular wagging run (e.g., the first, the fourth, or all of them) refers to a point of the search area functionally related to the goal (such as: that wagging run always refers to the boundary point which is due north of the goal), or by giving two boundary points lying on a diameter of the search area. If the first of these allows the rest of the SD to wander randomly over the search area, it poses a practical communication problem since potential recruits do not all follow an SD from the dancer's first cycle; if it requires that all (or all left circuit or right circuit) wagging runs refer to the specified point, it is an essentially uninteresting variant of the refer-to-the-goal syntactic strategy of the previous paragraph. As will be shown, a syntactic coding which gives the two boundary points directionally most distant from each other, i.e. along that diameter orthogonal to the hive-to-goal ray,  $r_0$ , in capturing the centrally important notion of the angular deviation, provides an explanation for all of the phenomena noted at the beginning of this section.

A first approximation to the principles of syntactic coding is given by (87), which identifies the pairs of points

$$\begin{aligned}
 (87) \quad & \{(r_1, \theta_1), (r_1, \theta_2) / \sqrt{(r_1 \cos \theta_1 - r_1 \cos \theta_2)^2 + (r_1 \sin \theta_1 - r_1 \sin \theta_2)^2} \} \\
 & = c, r_1 > 0, 0^\circ \leq \theta_1 < 360^\circ, 0^\circ \leq \theta_2 < 360^\circ \} \\
 & = c, r_1 > 0, 0^\circ \leq \theta_1 < 360^\circ, 0^\circ \leq \theta_2 < 360^\circ \}
 \end{aligned}$$

which are the requisite angular distance of  $c$  units apart. All three conditions ensure that the  $c$ -distant points are reported in a canonical notation; the  $r_1 > 0$  condition additionally characterizes the origin (hive) as a non-codable point, which of course follows from the general hive-to-point coding of the dance language. (87) is not as yet complete. It must be supplemented by an algorithm specifying which of the two points  $P_1 = (r_1, \theta_1)$  or  $P_2 = (r_1, \theta_2)$  is to the left and which is to the right of the goal  $P_0 = (r_0, \theta_0)$  (see (83)), for this is reflected syntactically by the +/- orientation of the return run it is followed by (i.e., L features a counterclockwise, R a clockwise return

run, and \*  ). One problem is immediately apparent--that of  $c$ -distant points satisfying (87), a full  $180^\circ$  apart. Neither  $P_1$  nor  $P_2$  is in that case the leftward point.

This is the first indication that the present approach is proper, for it is in precisely this  $180^\circ$  case that  $r_0 = 0$  and the coded goal is the hive itself. This view of the syntactic coding of the search area thus suggests that it is linguistically impossible to code the hive as goal, giving a different, not communication-based as was the round "dance" explanation for goals  $(r, \theta)$  with  $0 < r \leq c/2$ , account for this situation. Revising (87) as per (88)<sup>67</sup> allows for a quite straightforward

(88)

$$\left\{ (r_1, \theta_1), (r_1, \theta_2) / \sqrt{(r_1 \cos \theta_1 - r_1 \cos \theta_2)^2 + (r_1 \sin \theta_1 - r_1 \sin \theta_2)^2} \right. \\ = c, r_1 > 0, 0^\circ \leq \theta_1 < 360^\circ, 0^\circ \leq \theta_2 < 360^\circ, \\ \left. \theta_2 \neq \theta_1 + 360^\circ \ 180^\circ \right\}$$

mathematical reconstruction of the leftward property: whichever angle is in clockwise motion nearest the goal  $P_0$  is the leftward point (at  $\theta_L$  in (89))<sup>68</sup> of the pair. (88)

$$(89) \quad x, \bar{x} \text{ are such that: } \theta_1 + x = \theta_0 \quad \theta_2 + \bar{x} = \theta_0 \\ \min(x, \bar{x}) = x \implies \theta_1 = \theta_L, \quad \theta_2 = \theta_R \quad 0^\circ \leq \theta_R < 360^\circ \\ \min(x, \bar{x}) = \bar{x} \implies \theta_1 = \theta_R, \quad \theta_2 = \theta_L \quad 0^\circ \leq \theta_L < 360^\circ$$

must now be adjusted to include the return run C, oriented counterclockwise (+) following  $\theta_L$ , clockwise (-) following  $\theta_R$ , identified as in (89). The result is (90). In (90), SD is defined as n identical transitional dances of the

$$(90) \quad \left\{ [(r_1, \theta_R) C_- (r_1, \theta_L) C_+]^n, [(r_1, \theta_L) C_+ (r_1, \theta_R) C_-]^n / \right. \\ \left. \sqrt{(r_1 \cos \theta_L - r_1 \cos \theta_R)^2 + (r_1 \sin \theta_L - r_1 \sin \theta_R)^2} \right\} = c, \\ n \geq 1, r_1 > 0, \theta_R \neq \theta_L + 180^\circ, 0^\circ \leq \theta_L < 360^\circ, \\ 0^\circ \leq \theta_R < 360^\circ \}$$

form LR, or n of the form RL,  $n \geq 1$ . As will be shown, the full range of observed "dance" behavior can be explained on this model.

All that remains to complete the analog grammar, which does not generate but instead defines the utterances of the language, is to provide some means of wedding the

syntactic representations in (90) to the semantic representations in (84). This is achieved, simply enough, by a function taking as input the points  $P_1$  and  $P_2$ , and giving as output  $P_0$ , as the key to the semantic representation;  $P_0$  is the midpoint of the segment  $\overline{P_1P_2}$ , and is therefore given by (91).<sup>69</sup>

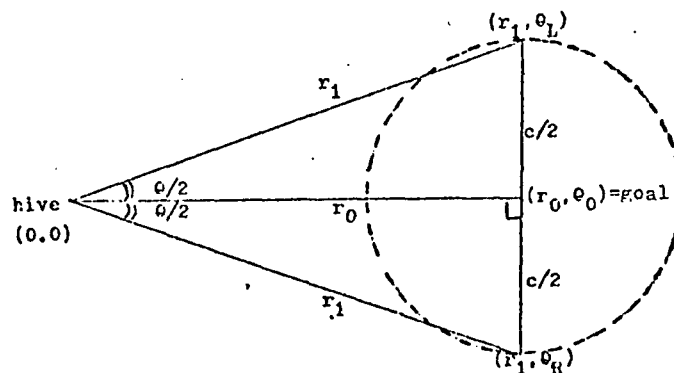
$$(91) \quad r_0 = \frac{r_1}{2} \sqrt{2 + 2(\cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2)}$$

$$\theta_0 = \cos^{-1} \left[ \frac{\cos\theta_1 + \cos\theta_2}{\sqrt{2 + 2(\cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2)}} \right]$$

$$\theta_0 = \sin^{-1} \left[ \frac{\sin\theta_1 + \sin\theta_2}{\sqrt{2 + 2(\cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2)}} \right]$$

This completes the analog grammar and thus the hypothesis as to what constitutes dance language competence. The goal  $P_0$  is the center of a search area of radius  $c/2$ , syntactically coded by reference to  $P_L$  and  $P_R$  in (92),

(92)



via wagging runs of equal duration (supposing wagging duration to be the linguistic coding of this variable), deviating by amount  $\theta$  in angular specification, where  $\theta$  is given in (93).<sup>70</sup> It has been noted repeatedly in the

$$(93) \quad \cos\theta = \frac{2r_1^2 - c^2}{2r_1^2}$$

dance language literature that  $\theta$  maximizes at  $60^\circ$ , which substitutes into (94) (derived by reference to (92)) to

$$(94) \quad \tan\left(\frac{\theta}{2}\right) = \frac{c/2}{r_0}$$

give a value of  $\sqrt{3} c/2$  for  $r_0$ . In other words,  $\sqrt{3} c/2$  is (roughly) the smallest goal distance, since  $60^\circ$  is the largest SD angular deviation (see (82)), for which there is distance coding.

It is of crucial importance at this time to determine what accounts for this  $\sqrt{3} c/2$  lower limit on  $r_0$ . Three avenues of explanation must be explored: the grammar, the production mechanism, and the comprehension mechanism.

The grammar has thus far had the responsibility solely for precluding directional coding of a goal at the hive or (if the R-constraint of (84) is adopted) beyond the flight range. The  $\sqrt{3} c/2$  limit on  $r_0$  certainly does not follow logically from either the semantic or syntactic coding system; a brute force introduction of

such a limit into the grammar would clearly be ad hoc, and so a fairly reliable indication of mis-analysis.

The point-to-patch cognitive net notion of production has thus far served to explain the round "dance" range of use; whenever the  $c/2$ -radius search area includes the hive, dance behavior is subject to an interference effect. At very small values of  $r_0$ , viz.,  $r_0 \leq c/2$ , the  $c/2$  distance variation at the semantic level may be expected to disrupt the syntactic form of the dance, giving a recognizable pattern. This production theory logically has only one other consequence: a disruption of dance form at very large values of  $r_0$ , where the  $\theta$  value becomes impossibly small, i.e. too small for that particular dancer to be able to perceptually distinguish the coding of  $P_L$  from that of  $P_R$ . Furthermore, this should by rights be a production problem with a greater variation across individuals, since it is so intimately tied to perceptual acuity.

The following quote suggests that in fact this is the case: "[c]ounts made on photographs showed S-curves only after 3-4 percent of the wagging runs for food places as much as 1000 m distant . . . , after 14 percent for 2000 m . . . , and 33.3 percent for 4500 m" (von Frisch 1967a:59). Surely not all this variation is legitimately tied to wind or gradient effects, with fully 14% of the 2000m flights more fatiguing (see sec. 2.4.1) than 66.7%

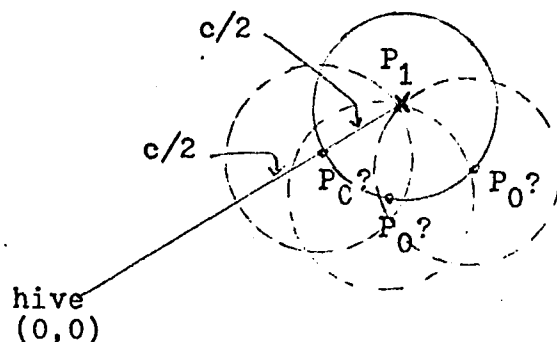
of the 4500m flights. The perceptual acuity demanded by the three distances of 1000m, 2000m, 4500m can be calculated via (94); since von Frisch's work was with the Carniolan race, which uses the round "dance" for a 16 m range,  $c = 32$  m, giving  $\theta$  values (see (92)) of  $1^{\circ}50'$  at 1000 m,  $0^{\circ}55'$  at 2000 m,  $0^{\circ}24'$  at 4500 m. The S-dance figures suddenly seem quite plausible.

The slight but steady increase in SD distance variation at such large  $r_0$  values can then be read as attempts to emphasize the search area semantic interpretation in cases where the  $P_L, P_R$  syntactic coding is not possible for perceptual reasons. The cognitive net theory of production has no grounds for supposing that directional coding is in any way problematic for goals between  $c/2$  and  $\sqrt{3} c/2$  units from the hive. There is no way for a production mechanism of the kind outlined to predict more than one small- $r_0$  execution problem, must less fix values for more than one limit. The explanation for the  $\sqrt{3} c/2$  limitation must be derived from a theory of SD comprehension.

The simplest comprehension theory is that which most economically relates the processing in comprehension to that in production. The averaging semantic rule of (94), which finds  $P_0$  by bisecting the segment  $\overline{P_1 P_2}$  bears no discernible resemblance to the cognitive net notion of production. The point-to-patch process, however, works

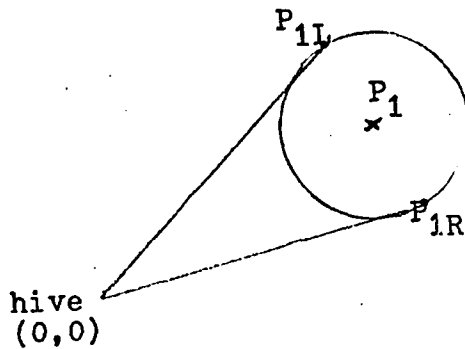
equally well as a comprehension model--better, in fact, since there is no need for a codicil discriminating between the left and right danced values  $P_1$  and  $P_2$ . The mechanism is simply this: since  $P_L/P_R$ , the syntactically coded values, are exactly  $c/2$  units from  $P_0$ , a point-to-patch  $c/2$  extension from a received  $(r_1, \theta_1)$  must include, on its boundary, the goal  $(r_0, \theta_0)$ . The  $c/2$ -sized search area is thus known to be centered on one of the boundary points (see (95)).

(95)



Now, the canonical transitional form of an SD is such that a comprehension mechanism could be proposed which would fully determine the position of the goal,  $P_0$ , from that single point,  $P_1$ . That strategy is as follows: (a) read off from the  $c/2$  point-to-patch extension about  $P_1$  the two boundary points ( $P_{1L}$  and  $P_{1R}$  in (96)) which are directionally most distant--one of these is  $P_0$ ; (b) if the orientation of the curve following the  $(r_1, \theta_1)$  wagging run is counter-clockwise (+),  $P_1 = P_L$  and  $P_0 = P_{1R}$ , while if the orientation is clockwise (-),  $P_1 = P_R$  and  $P_0 = P_{1L}$ .

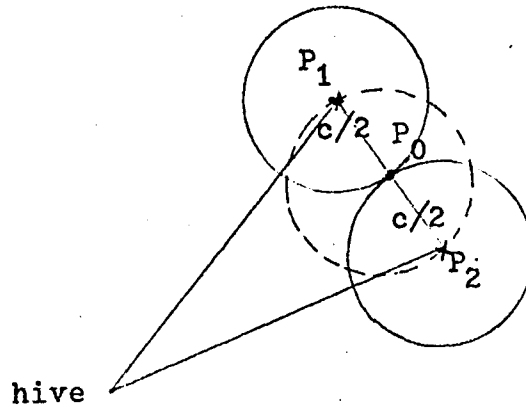
(96)



The problem with this comprehension strategy is that it depends on the observer's accessing all of the information present in the signal, down to the curved run, which the available evidence on dance following suggests is not very significant in communication. What's more, the 6-cycle perceptual averaging of sec. 2 becomes either a prohibitively complex 3-cycle  $P_L$ , 3-cycle  $P_R$  averaging or else is an averaging of 6 estimates of  $P_0$ , which since  $P_0$  must be calculated from the physical signal, is not obviously perceptual in nature.

According to the simpler comprehension model first developed (see (95)),  $P_0$  is not calculated (although as just outlined it is calculable) from a single circuit. The second wagging run is interpreted as was the first, i.e. as a point  $P_2$  about which a  $c/2$  sized patch is drawn; the patches intersect at  $P_0$  (see (97)). "Averaging" is not so much averaging to  $P_0$  as it is targetting  $P_0$  by an intersection method utilizing circles of radius  $c/2$  (rather

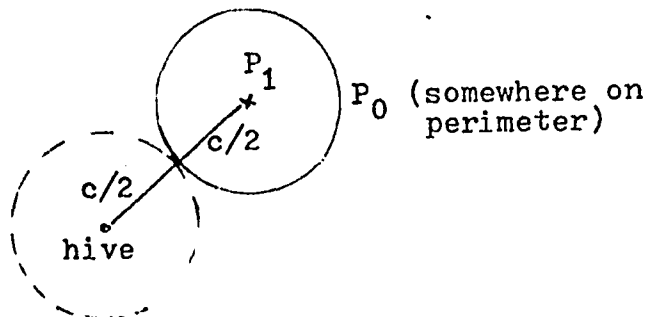
(97)



like the Ballantine trademark).  $P_0$  is ideally part of every patch's perimeter; in practice, merely common to most of them. Unlike the comprehension model associated with (96), this works for the S-"dance" behavior as well. What's more, again unlike the previous comprehension model, this one<sup>71</sup> explains the  $60^\circ$  maximum on angular deviation.


According to the Ballantine model, any executed wagging run is extended from the coded point to a  $c/2$ -sized patch on whose perimeter the goal  $P_0$  may be found. Now, as already discussed,  $P_0$  is minimally  $c/2$  units from the hive (lest the search area include the hive).  $P_1$  is thus minimally  $c/2 + c/2 = c$  (see (98)) units from the hive.

(98)



Then the angular deviation  $\theta$  (see (92)) is maximally (see (82)) such that  $\cos\theta=1/2$  (using (93), with  $r_1=c$ ), i.e.,  $\theta$  is maximally  $60^\circ$ . The same restriction can be stated in terms of  $r_0$ , which is minimally  $\sqrt{3} c/2$  (by the Pythagorean theorem). The  $60^\circ$  limit has a principled basis. In fact, so does the shape of the behavior evinced in such performance-plagued circumstances.

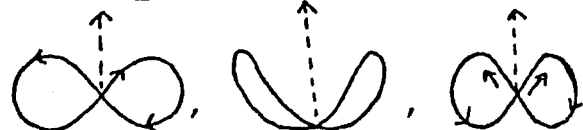
The dance language is quite simply more iconic than has been suspected; for  $r \leq c/2$ , where the problem is one

of production, the "dance" is a simple circling pattern , while for  $c/2 < r \leq \sqrt{3} c/2$ , where the problem is one of comprehension, due to the dancer's simultaneous functioning as transmitter and receiver,<sup>72</sup> the "dance" has two circles





. The same point may be made with respect to the transitional and tailwagging dances. These too are far more iconic than is generally realized; their twin semicircle construction recapitulates the shape of the search area, halved by  $r_0$  into points leftward vs. points rightward of the goal relative to the hive.

Neither is it surprising that for this problematic range ( $c/2 < r \leq \sqrt{3} c/2$ ), direction though not distance is

given: . The situation arises, after all, due to a comprehension problem caused

by the distance variation's magnitude of  $c/2$ . That is, it is the potential  $P_0^{73}$  along the ray  $r_1$  (see (98)) which sets the limit; that  $P_0$  differs from  $P_1$  only in distance, not in direction.

Finally, it makes sense that the form the gradual

transition from  to  takes is a matter of racial variation and genetic coding. The "dances" in this  $r$ -range ( $c/2 < r \leq \sqrt{3} c/2$ ) are the only ones whose shape is not predictable from the search area hypothesis. It is thus the only possible locus for genetic variation, the only possible source of "dialects," these being due to variation in innately specified responses to a performance problem, not different grammatical systems.

The search area hypothesis, construed in these syntactic terms, with a search area centered at  $P_0$  coded by reference to  $P_L$  and  $P_R$ , treats directional but not distance variation as part of the syntactic system. It is clear that the descriptive and "explanatory" advantages of this approach depend on syntactic dignity for directional variation. In particular, no account can be given for the observed diversity of dance types, their distribution, or the parameters of genetic variation unless directional variation across SD is syntactic. What is not clear is the status of the argument with respect to distance variation across SD.

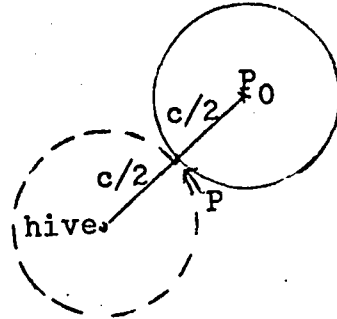
The search area hypothesis in its present form essentially holds that distance and directional variation are independent--and theoretically separable along a competence/performance dimension, this separation linguistically motivated. It will now be shown in what respects a syntactic analysis treating these as dependently variable fails to account for the phenomena. The analysis in question treats SD as a seriated specification of within-the-search-area points, selected virtually at random alternately from the left and right search area halves.

From the point of view of this approach, every issue is one of statistical confidence. Consider, for example, the problem of explaining the  $60^\circ$  upper limit on  $\theta$ , the  $\sqrt{3} c/2$  lower limit on  $r_0$ . In this analysis, it reduces to a problem of determining how far away (hopefully, some approximation to  $\sqrt{3} c/2$ ) a goal must be from the hive in order to guarantee that 6-cycle (see p. 65) averages of random points at most  $c/2$  from that goal will have some independently motivated minimal probability of being at least  $c/2$  away from the hive, and thus outside the "danger zone," this being the area in which search area construction has a sufficiently high probability of including the hive.

Now, whatever this minimal value ( $r_{\min}$ ) is, it is smaller than  $c$ , since for  $r_0 = c$ , there is a probability of virtually 1.00 that 6-cycle averages are outside the

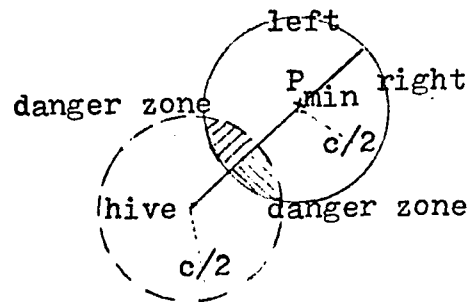
danger zone; it would take 6 instances of P in (99) to

(99)



convey a hive-included search area. If what is required is statistical confidence on the order of, say .95, then some degree of overlap (see (100)) is to be tolerated between the danger zone and the search area about  $P_0$  defining

(100)



the domain of points susceptible to the select-one-at-random SD construction process. How much overlap (which might be given, for example, in terms of the percentage of area which can be shared in that way) is a function of this independently motivated, genetically coded confidence factor;  $r_{\min}$  is then a function of the degree of overlap which can be tolerated in this way. It will not be easy to motivate genetically coded confidence levels that discriminate between the cases provoking a round dance and

those provoking the direction-coding nonwagging transitional dances. In fact, (100) is clearly an oversimplification, since it assumes that the partially intersected patches have the same radius. This is to say that the round dance confidence level is 1.00, which of course is suspiciously high. Were all the confidence levels that high, the  $P_R$ ,  $P_L$  idealization of the earlier syntactic strategy would be supported.

In conclusion, an analysis syntactically dignifying SD directional variation while treating distance variation as noise engendered by the psychological reality of the search area hypothesis at the semantic level, is to be preferred.

The dance language is a rule-guided system, extending the site of a stimulus to a patch of uniform size, syntactically coded via reference to the directionally most extreme points of the patch. It is far more iconic than has previously been supposed, yet includes a cognitive, competence level.

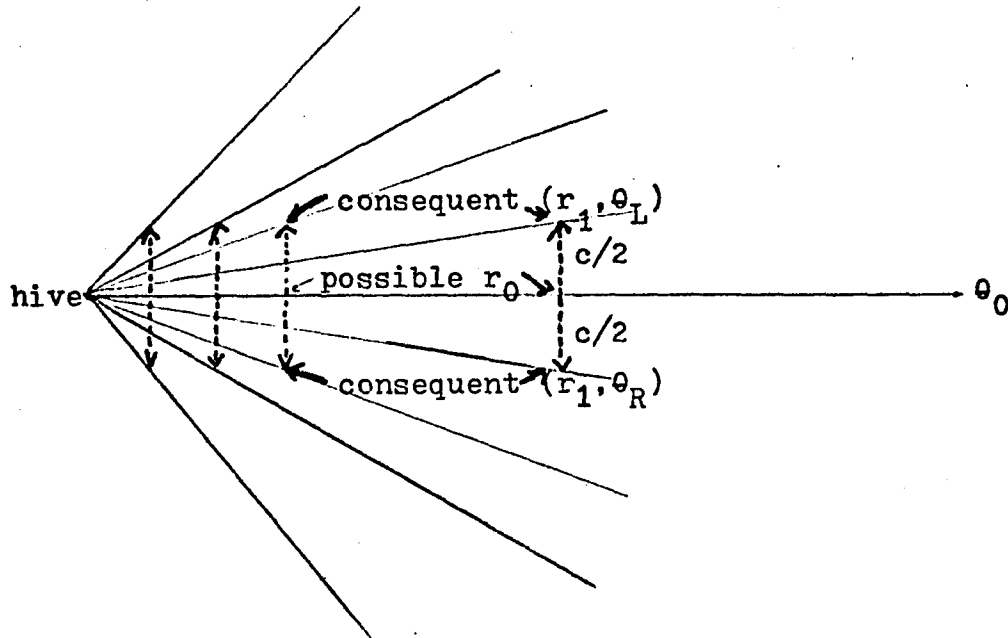
Gould (1976:231) complains that "to be consistent, the teleological theory must assign a plausible purpose to the details of the dance." The quest here for teleological "explanatory" adequacy is just such a quest; the search area hypothesis does just that. Within the context of this grammatical theory, there is a natural explanation for the observed variety of dance types, their hierarchical

ordering vis à vis the distance from the hive to the goal, the very limited degree of genetic variation, the fact that wagging runs are at most  $60^\circ$  apart. The search area hypothesis is at once observationally, descriptively, and "explanatorily" adequate.

Now, in terms of the linguistic evaluation of the dance language, the most important aspect of the search area hypothesis, properly construed, is this. The vocabulary is not nondenumerable solely because there are nondenumerably many goals to be encoded, but because the message sent for any particular goal is of this  $c$ -distant  $P_L, P_R$  kind. The search area hypothesis actually entails that the dance language is not merely syntactically different in degree from human language. That is, there is no cognitively principled (see sec. 4.2) countable set which fits this syntactic strategy.

What is required of the lexical set is that for any lexical specification  $(r_0, \theta_0)$ , there be two distinct lexically specified points  $(r_1, \theta_1)$  and  $(r_1, \theta_2)$  each an angular distance of  $c/2$  from  $(r_0, \theta_0)$ . That is, for any lexically coded ray  $\theta_0$ , each lexically codable  $r_0$  must have a  $c/2$  distant  $(r_1, \theta_L)$ , forcing the construction of a set of rays, one for each  $r_0$  candidate; the same applies for  $(r_1, \theta_R)$ . The situation may be pictured as in (101). Any  $r$ -values thereby lexically paired with a  $\theta_1$  must also be lexically paired with all other  $\theta$ 's, including for example

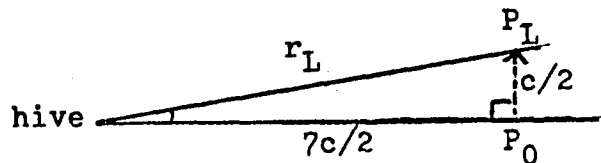
(101)



$\theta_0$  itself.

A syntactic strategy which proceeds using a  $c/2$  measure requires a syntactic coding of  $c/2$ , and thus of  $7c/2$ --since  $7c/2$  is certainly not greater than the foraging range ( $\theta$  would be approximately  $16^\circ$  leaving it well within the "normal" transitional dance range) and the lexicon must be cognitively principled. Consider the lexical specification  $(7c/2, \theta_0)$ ; by the Pythagorean theorem (see (102)),

(102)



$r_L = 5\sqrt{2} (c/2)$ , an irrational multiple of the lexically

coded  $c/2$ . A cognitively principled countable lexical set can not include specifications that are irrational multiples of each other.

The "explanatorily" adequate construal of the search area hypothesis requires a nondenumerable  $\omega$ -range and thus demonstrates that the dance language is a rule-guided system which is syntactically different in kind rather than degree from human language.

## Notes

4

## ALTERNATIVE GRAMMATICAL ACCOUNTS

1. Herman and Rozenberg 1975 proposes a different hierarchy of grammars, recognizing more than four alternatives. The traditional four-way distinction in (36), however, has the merit of being well-investigated vis à vis (human) linguistic theory, and so was to be preferred in this context.

2. This strategy was perhaps most obvious in the case of the round dance (see pp. 120-22), but operated throughout sec. 3. Thus, since the goal's location is discoverable only via averaging across the message series (see p. 65), the transitional dance is generated in, for example, (14), as a left and a right circuit which average to the proper direction. This incidentally has the additional (syntactic, this time) merit of imposing on the transitional dance the both-right-and-left-circuits property of the superdance. It is largely for this syntactic reason that the 2-circuit analysis for the tailwagging dance is pursued in (12), although the concomitant structural similarity to the 2-circuit transitional dance is certainly desirable. Finally, the S-shaped curved runs that serve to separate the wagging runs during an S-dance series are treated in sec. 3 as each a part of the previous wagging-run-introduced dance.

3. If supported, it should be written as  $D \rightarrow DD^*$ . The SD symbol is retained to keep the discussion as clear as possible.

4. For some specific values of  $i$ , however, it may be. Thus, for  $i$  very large,  $\omega_i \propto j$  may be appropriate only to an S-dance, allowing for message interpretation immediately upon the perception of  $\omega_i \propto j$ . If, of course, the  $\omega$ -grid is so fine that each  $\omega$ -value is appropriate to only one dance type and those tied to the transitional dance form are appropriate to only one particular degree of deviation from the bisector (so: a left circuit  $\omega_i \propto j \oplus 5$   $c_{i,j} \oplus 5, +$  is only used for a transitional dance with a  $2-\omega$  value deflection from its bisector, and so can only be completed by a right circuit  $\omega_i \propto j \oplus 9$   $c_{i,j} \oplus 9, -$ ) then message recovery is

assured from the first wagging run, curved run combination. (The curve is needed to ensure that  $\omega_i \propto j \oplus 5$  appropriate only to a 2- $\omega$  deflection, is the left circuit component;  $\omega_i \propto j \oplus 5$   $c_{i,j} \oplus 5,-$  would presumably be completed by  $\omega_i \propto j \oplus 1$   $c_{i,j} \oplus 1,+$  with the message being not  $\omega_i \propto j \oplus 7$  but rather  $\omega_i \propto j \oplus 3$ . It would take an additional grammatical constraint ensuring either that left or that right circuits were uniformly initial in transitional dances, to permit message recovery from the wagging run alone.)

5. Note that no proof has been given that rules of higher type, as for example context sensitive phrase structure rules, absolutely can not account for the incrementing.

6. Consequently, as was pointed out above, its structural status (see n.5) is of little moment and will not be pursued here.

7. So, based on afternoon daylight experiences, bees raised in a dark cellar reconstruct the morning portion of the curve (Lindauer 1959, as cited in Lindauer 1967:121-26). As is also pointed out in Lindauer 1967 (pp. 123-24, 95), the solar incrementing ability displayed during after-sunset site advertisements which are often part of marathon dancing similarly indicates ability in solar transit extrapolation.

8. Lindauer 1967:118-19 reports conflicting experimental results in this regard. Kalmus 1956 (as cited in Lindauer 1967) claims that at least some aspects (such as, perhaps, clockwise vs. counterclockwise diurnal transit) of solar motion are innate, while Lindauer's own work, although it argues for a simple conditioning continuation extrapolation account of solar motion learning, is itself equivocal with respect to whether conditioning can cause an earlier acquired system to be substantially revised.

9. Undesirable unless, of course, only inexperienced bees in the language acquisition (pragmatics acquisition?) stage display disoriented dancing in this near-zenith situation.

10. Necessarily so, if solar motion is to be deduced from it.

11. This is not to say that bees are utterly unable to perceive the magnetic field. On the contrary, the magnetic field aids in orienting the comb in new hives as it was in the parent hive, and is additionally responsible for vertical dance floor misdirection (Gould 1976:215, reporting results of Lindauer 1973, 1974; Martin and Lindauer:

1973, in the first case, and Lindauer and Martin 1968, in the second). What is being suggested is that solar motion is not learned by reference to this factor; the hypothesis would be falsified if varying the magnetic field diurnally with solar motion leads to cellar-bee behavior. The discovery that solar motion is learned by reference to the magnetic field would be particularly interesting, for it would suggest that magnetic field in-flight orientation is not possible--perhaps because the visual stimulus of the sun is biologically so obtrusive, perhaps because the flight mechanism somehow blocks magnetic field perception, perhaps because olfactory searching precludes magnetic field perception.

12. Reliable experimental evidence that only unbroken linear displays compete with the solar compass with respect to orientation in flight would cast doubt on the A-analysis since it would suggest that A must be of that linear nature. Then, since not all hive locations would have a linear landmark in the vicinity, the theory of solar incrementing acquisition would become suspiciously complex.

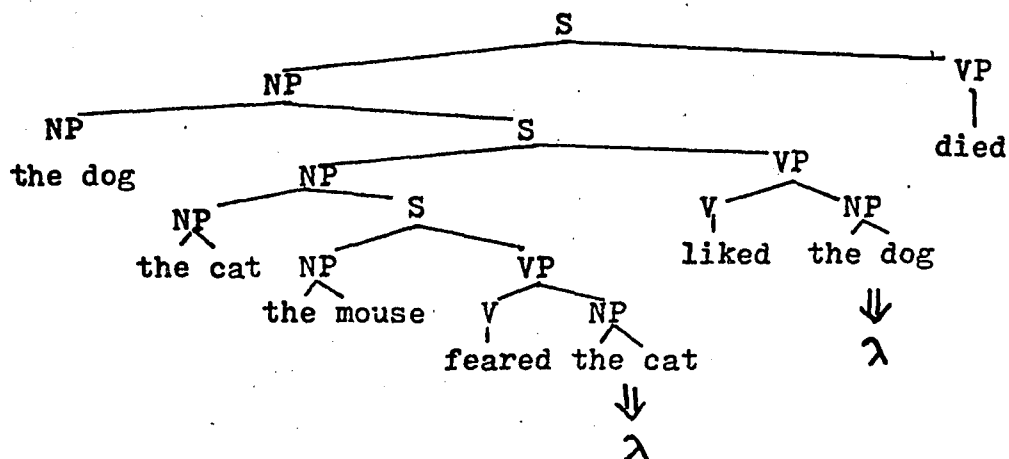
13. As will be shown, there is some evidence of the need for context sensitive but none for transformational capacity with regard to the goal of observational adequacy.

14. The rest of this section is drawn largely from Levelt 1974b. Also, see Langendoen 1976 for a critical discussion of the relationship between these assumptions about grammaticality and the notion of grammar as an account of the internalized product of the language acquisition process.

15. See Schlesinger 1968 for additional examples and discussion.

16. If, on the other hand, (55) is grammatical but semantically anomalous, so is any string associated with a completion of (56). In either case, the general point of projecting a linguistic property of (56) from those of (55) stands.

17. The usual analysis makes all grammatical relations, such as subject and object, explicit at the sentence level, transformationally deleting the second co-referential occurrence of an NP whose sentence is in construction with the first occurrence, as here:



18. This is actually a much stronger assumption than need be made; it was not after all necessary to identify all of English with  $\{(the\ oyster)^n\ (split)^n / n \geq 1\}$ . What must be minimally guaranteed is that all elements of (57) are grammatical and additionally that there are no grammatical series having all left circuits first but a larger (or smaller) number of subsequent right circuits. We ignore entirely the SD version identical to this save for generating right circuits initially.

19. But see Postal 1963's argument that Mohawk is not context free.

20. The implications of this observation are further examined in sec. 4.1.3 and sec. 4.3.

21. What has been behaviorally supported is a two-feature decomposition of the wagging run. The particular nature of that decomposition is further discussed in sec. 4.2, sec. 4.4.

22. There is very little in the literature dealing specifically with the curved run and therefore no attempt is made to behaviorally justify the full terminal string.

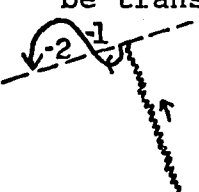
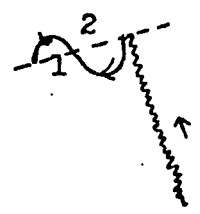
23. Since no evidence has been presented for a behavioral category curved run (or clockwise curved run or counterclockwise curved run either, for that matter), the symbols  $c_{i,j,+}$  and  $c_{i,j,-}$  have no more status here than a deictic or referential indicator.

24. Of the sixteen design features outlined in Hockett and Altmann 1968, only six can be understood as grammatical desiderata: duality of patterning, discussed here and discrete vs. continuous signal repertoire, discussed in sec. 4.2; openness, arbitrariness, semanticity, and reflexiveness are discussed, along with the other design features, in sec. 4.3.

25. See for example, Thorpe 1974:70, Hockett 1960: 428.

26.  $(\omega_{I_1}, \omega_{I_2})$  is the range of circuit-appropriate  $\omega$ -values.

27. An S-dance of the form  could conceivably

be transformationally related either to  or to 

In the first case, a R-L metathesis rule is involved, thus requiring evidence for an L vs. R grammatical

distinction; in the second case, a  $\begin{Bmatrix} + \\ - \end{Bmatrix} \Rightarrow \begin{Bmatrix} - \\ + \end{Bmatrix}$  rule is needed to slip the curves about their (non-executed--) axis, requiring evidence for a + vs. - grammatical distinction.

28. A rule of the form  $LR \rightarrow AR$  is often written  $L \rightarrow A/_R$ , i.e. "L" goes to "A" in the context "before R."

29. Simplicity plays such a stellar role here due to the assumption that transformational power has been established for the dance language via a demonstration of the grammatical need for spread rules. Thus the notion of behavioral evidence for metathesis (+/- switching) rules plays a much less vital part in motivating the rules than it did earlier.

30. For example,  $80^\circ$  and  $100^\circ$ ,  $140^\circ$  and  $160^\circ$ .

31. For example,  $330^\circ$  and  $280^\circ$ ,  $200^\circ$  and  $180^\circ$ .

32. Although this has been treated as extragrammatical, it has been argued (see discussion, p. 177) that this is a defect of the grammar thus far proposed.

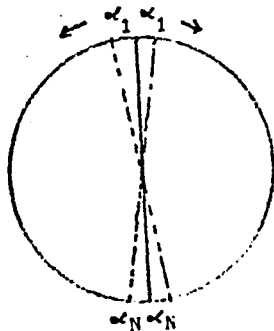
33. See, for example, Gould 1975c:170; Gould 1976:239.

34. See the search hypothesis outlined in Gould 1975b: 173-76.

35. The single exception is discussed in sec. 4.3 with regard to Hockett's design feature of displacement.

36. See for example, Marler 1970; Nottebohm 1968, 1971, 1972, 1975, and Nottebohm and Nottebohm 1976.

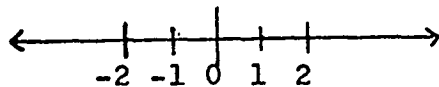
37. Von Frisch's position is obviously closer to



than (76). However, for  $j$  sufficiently equidistant between 1 and  $n$ , the sequence  $\omega_i \propto_j c_{i,j,+}$   $\omega_i \propto_j c_{i,j,-}$  is only possible if the  $j$ 's are either all left- $j$ 's or all right- $j$ 's; otherwise the direction is being given by wagging runs more than  $60^\circ$  apart. To avoid this obvious problem, von Frisch's position is given the more charitable interpretation of (76).

38. See discussion, pp. 188-91.

39. A word of explanation as to the mathematics involved in this discussion. An infinite set is one which can be put into one-to-one correspondence with a proper subset of itself (intuitively, a set which can lose some of its elements without getting smaller). Any set which can be put into one-to-one correspondence with the set of natural numbers,  $\{1,2,3,\dots\}$ , is said to be denumerably infinite. Examples include  $\{2,4,6,\dots\}$  and the rational numbers, i.e. the set of fractions. The real numbers (those which can be located as a point somewhere on the number line



are provably nondenumerably

infinite as is any bounded set of real numbers (e.g. those between 0 and 1). There are, then, more real numbers than natural numbers but not more natural numbers than even numbers. A grammar  $G$  for a language  $L$  may be viewed as an

enumeration 1,2,3, ... of the sentences of L; it is therefore not possible for any grammar G to generate a nondenumerable set of sentences. Now if there are nondenumerably many  $\omega$ -values, then there are nondenumerably many dances to be accounted for.

40. The intrinsically structural criticism of (9) is discussed here rather than in sec. 4.1 for heuristic reasons.

41. Always supposing that the failure of observed SD's (or, for that matter, full-wagging transitional dances) to code an integral number of wagging movements (or sound bursts) is not best explained as an execution problem, in which case the  $\omega$ -grid may be given directly in those terms, with m the maximal number of wagging movements (or sound bursts) produced.

42. Regardless of the relationship of M and p, a grid of size  $m = Mp$  will allow for integral message coding. If  $q > 1$  is the maximal integral factor common to p and M then the smaller value of  $Mp/q$  will do for m.

43. Quine 1960:26-40 discusses the behaviorist pitfalls encountered in determining the translation of a term "Gavagai" from a previously unencountered human language to English, arguing that there is no "best"--and so, no correct--translation.

44. It is an open question whether in this event it would prove to be experimentally undecidable and thus a genuine Gavagai case.

45. The assumption throughout this discussion of possible grammatical accounts is that there are only finitely many  $\omega$ -values, since of course no grammar may be written for a nondenumerable  $\omega$ -range lexically based language.

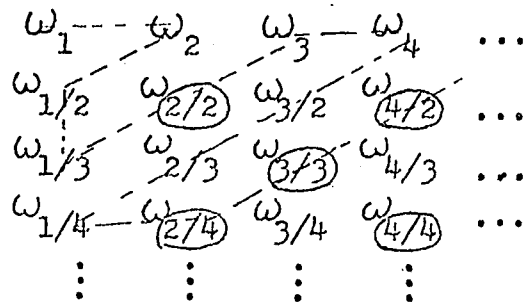
46.  $\omega_0$  is the smallest wagging duration value, which may not correspond to the constant  $\omega_1$  differential between adjacent  $\omega$ -values.

47. The anti-M criticism is motivated by the distance semantics traditionally assumed and the concomitant bias toward a competence idealization allowing arbitrarily large distance values to be encoded. Since the round dance is restricted in use to a local distance range it is not in the sense of the anti-M objection to grammatically vary the round dance in a denumerable way (as by extending the one circle, one reversal pattern of (26), viz.  $c_{i,j,+}$   $c_{i,j \oplus n/2,+}$   $c_{i,j,-}$  and  $c_{i,j,+}$   $c_{i,j \oplus n/2,-}$   $c_{i,j,-}$  to the full range of

i values). It would admittedly not be within the power of the finite state device to capture the i-repetition dependence for a denumerable range of i's. However, it would take more than the anti-M position to indicate that the round dance is so variable.

48. For the sake of uniformity, the finite state account of the other dance forms will also use this revised notation  $\omega_{k+1} \propto_j$  being generated as the  $k + 2$  sequence  $\omega_0 (\omega_1)_k \propto_j$ .

49. The function attacks a rational display of the following kind at the upper left (grounded) corner and zig-zags, as drawn, through the table, assigning a natural number to each new value (circled values are ignored since not



new) in the order in which it is encountered. The full display is needed for anti-M anti-p; for anti-p, each (horizontal) line of the table will terminate when the subscript threatens to exceed M. (In the anti-p case, the values might of course alternatively be counted across, row by horizontal row, since each row will be only finitely long.)

50. With regard to the second of these, specialization--the requirement that the signal have only a communicative function--it has been shown (see discussion, sec. 2.13) that it does have a communicative function; although there is no way to prove this is its sole biological significance, no other has been proposed. It is generally said (see Thorpe 1974:70; Hockett 1960:428) that the dance language only partially meets the criterion of Interchangeability because only worker-bees, i.e. potential foragers, are linguistic participants. That criterion is, however, satisfied in that all linguistic participants can function as both transmitters and receivers.

51. As discussed in sec. 2.4.1, Gould 1975a:171 assumes that the production and comprehension mechanisms are equally efficient and outlines a statistical test to isolate the effective distance signal(s), subject to this assumption. Each distance correlation, and combination of correlations, is tested by the sum of the squares method (see

above, pp. 33-34) in a misdirection (see sec. 2.13) context. The candidate for which the recruit scatter is reliably approximated by  $1/\sqrt{2}$  the dancer's scatter is the distance signal. The assumption of equally efficient distance encoding and decoding mechanisms is simultaneously confirmed by the isolation of such a reliable approximation. It is further noted (again, Gould 1975a:171) that recent work suggests that the dance sound may be perceived via the antennae rather than the legs. Ultimately, the hypothesis must be tested by means of the model bee (see Gould 1975b:62-75 for a discussion of the progress that has been made along these lines), which remains the only way to vary wagging and sound production independently.

52. I am indebted to Jerrold Katz for this point.

53. Specifically (see Bennett 1964:12-15), that the utterances be about something (Semanticity, discussed below), that the same-species receiver comprehend the message (design features are for communication systems), that the language be creative (Openness, discussed below), and syntactically complex. It is noted in this regard that syntactic complexity may be met by simultaneous as well as by serial complexity; "There is no need to insist that such features must consist in the utterances' containing parts or episodes each of which has an independent linguistic significance" (Bennett 1964:13-14).

54. As will become obvious, regular is being used here in a technical sense different from the standard (a regular language is finite state generable) way.

55. This example taken from Gould 1975b:199.

56. See Chomsky 1975a, especially 3-35.

57. Limber 1977:282. Sec. 4.1.2 suggests that both of these ("sentence" is dance, D) may be descriptively motivated. Limber 1977, in contrasting chimpanzee and child linguistic behavior, only indicates the universals which discriminate between them. No very general claim is made.

58. However, as has been discussed, drawing a competence/performance distinction in the first place depends in a crucial way (it has been argued) on there being a degree of stimulus freedom to be uncovered. Appropriateness also remains important as a methodological issue, since it would be difficult if not impossible to linguistically analyze a system employed by a "schizophrenic" species.

59. And perhaps a third, phonological/cheremic/choreographic? parameter as well, which still remains to be generalized from human vocal language. Healy 1972 argues for a phonemic parameter which, since it presumes American Sign Language is not a natural language, is here ignored.

60. Since referential ambiguity precludes message recovery.

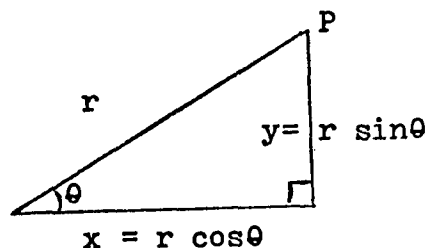
61. Later on that page, Chomsky specifically mentions the honey bee as an example of such a signaling system, qualifying his comment, however, to apply only under a favorable resolution to the dance language controversy.

62. Haldane and Spurway (1954:263) also use the speedometer example: "The bees' symbolism is like that of an analogical calculator such as a slide rule or a speedometer, which represents one continuously variable magnitude by another." Their use of the term "symbolism," however, suggests that their position may be closer to that of Janda 1977 than Chomsky 1975b. It is difficult to tell, inasmuch as no argument is given in support of their claim.

63. And which is questionable on either the flight range percentage or fatigue reading of the wagging run's duration.

64. That is,  $\omega_i$  is not the  $i^{\text{th}}$   $\omega$ -value, but some  $\omega$  value, with  $i$  functioning as an indicator of cross-reference possibilities.

65 This is also noted by Hockett (1960:412). Polar coordinates give the positioning of a point  $(r, \theta)$  in terms of its distance  $r$  and direction  $\theta$  from the origin. Throughout, points are specified in the canonical notation of non-negative  $r$  and  $\theta$  between  $0^\circ$  and  $360^\circ$ , with  $\theta=0^\circ$  denoting (by convention) the positive  $x$ -axis. The Cartesian coordinate system will give  $P$ 's coordinates as  $(x, y)$ , the polar coordinate system as  $(r, \theta)$ . A circle of radius  $c/2$  centered



at Cartesian coordinates  $(x_0, y_0)$  has the familiar equation:  
 $(x-x_0)^2 + (y-y_0)^2 = (c/2)^2$   
 Substitution of  $r \cos \theta$ ,  $r_0 \cos \theta_0$ ,  
 . . . for  $x$ ,  $x_0$ , . . . gives the  
 formula in (83).

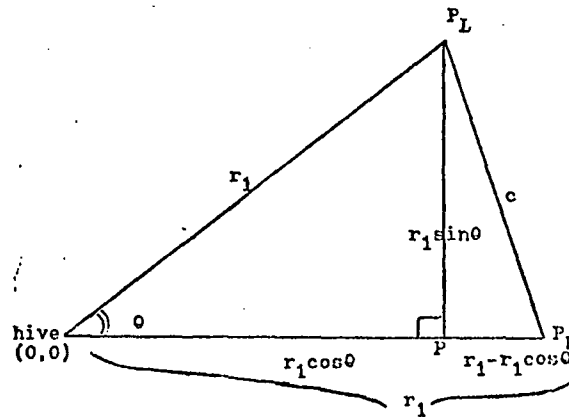
66. It is argued below (pp. 270-72), that the best version of the search area hypothesis entails a nondenumerable lexical vocabulary.

67.  $+360^\circ$  denotes addition modulo  $360^\circ$ , as per sec. 3 n.5.

68.  $x, \bar{x}$  are in the  $0^\circ$ - $360^\circ$  canonical form of n.65. Thus their minimum gives the leftward angle. Consider, for example,  $\theta_1 = 40^\circ, \theta_2 = 350^\circ$ ; then  $\theta_0 = 15^\circ, x = 335^\circ, \bar{x} = 25^\circ$ , and  $\min(335^\circ, 25^\circ) = 25^\circ = \bar{x}$ , identifying  $\theta_L$  with  $\theta_2$ , as desired.  $\theta_0$  is obtained from  $P_1, P_2$  via (91).

69. (91) was derived from the Cartesian counterparts  $x_0 = \frac{x_1+x_2}{2}, y_0 = \frac{y_1+y_2}{2}$ , using the rules  $x_0=r_0\cos\theta_0, y_0=r_0\sin\theta_0, x_1=r_1\cos\theta_1, y_1=r_1\sin\theta_1, x_2=r_1\cos\theta_2, y_2=r_1\sin\theta_2$ . (See n.65 for a discussion of the  $(x,y) \rightarrow (r,\theta)$  rules.)

70. Stripped of irrelevant detail, (92) is as follows:



Applying the Pythagorean theorem to  $\triangle P_L P P$  yields the formula  $c^2 = r_1^2 \sin^2 \theta + (r_1 - r_1 \cos \theta)^2$  which reduces to (93).

71. This ("Ballantine") model identifies  $P_0$  as the point  $c/2$  distant from  $P_1$  which is also  $c/2$  distant from  $P_2$ . The earlier model first identifies  $P_{1L}, P_{1R}$ , the two points an angular distance of  $c/2$  from  $P_1$ , and then if  $P_1 = P_L$

deduces that  $P_0=P_{1R}$ , and if  $P_1=P_R$  deduces that  $P_0=P_{1L}$ . It is possible to construct a third alternative, one which identifies the two points  $P_{1L}$ ,  $P_{1R}$  an angular distance of  $c/2$  from  $P_1$ , and, comparing these to  $P_2$ 's  $P_{2L}$ ,  $P_{2R}$ , picks out as  $P_0$  whichever point appears in both pairs. This analysis is a needlessly complicated version of the Ballantine model, requiring two wagging runs for  $P_0$ -determination but enjoining an additional processing step on the dance follower than is necessary. It additionally fails to predict the observed  $60^\circ$  maximum on angular deviation.

72. This incidentally further undermines the Bennett claim that the dance language is regular rather than rule-guided. "Reflexes" are not self-monitored for informational content in this way. What could be the transducer explanation for this  $60^\circ$  limit?

73. "Potential" from the point of view of what the performance process calculates, not what is calculable. A similar distinction is made in Katz 1977:571-73. The terminological tension there is between prediction and calculation which roughly corresponds to the calculated vs. calculable distinction drawn here.

## IMPLICATIONS, EXPERIMENTAL AND LINGUISTIC

The present inquiry into the structural properties of the dance language of the honey bee has the broader significance of illustrating a theoretically grounded methodology for the comparative linguistics of animal communication systems. In the past, this comparative issue has been addressed largely in an evaluative spirit (as, for example, in Hockett and Altmann 1968, or even Limber 1977), and while that certainly is of concern here, the emphasis has been shifted to a characterization of the nonverbal system in and of itself, as a physical signal-message correlation.

With the exception of the regular vs. rule-guided, sign vs. symbol issue, analyses of animal communication systems have typically been focussed entirely on properties of either the set of physical signals or the message set. Pankova 1972, Dethier 1957, Konishi 1970 and Bossert and Wilson 1963 all analyze the set of physical signals; Smith 1969, Katz 1976, Haldane and Spurway 1954 all deal with the message set. This linguistic (grammatical) investigation of the dance language is concerned with the mathematical properties of the system relating these sets

for the dance language.

The sign vs. symbol, regular vs. rule-guided debate finds Bennett 1964 of the opinion that the dance language is regular, Thorpe 1974 persuaded it is symbolic, and Kroeber 1952 suspecting that the directional indication with respect to gravity is symbolic, even as his acquisition criterion labels it non-symbolic. With regard to the various chimpanzee language projects, the sign/symbol problem has persisted as a major concern at the word level. (see, for example, Bronowski and Bellugi 1970 and the "wordness" hierarchy described in Rumbaugh and Savage-Rumbaugh 1978). For the honey bee, the issue arises at the sentence or discourse level, and is appropriately characterized in terms of the applicability of a competence/performance distinction to the dance language. As already noted, Chomsky 1975b claims, but offers no evidence or argumentation in support of that claim, that the dance language is syntactically different in kind rather than degree from human language because the dance language is regular rather than rule-guided, because, that is, there is only one analytical level, that of the physical signals themselves.

The question of whether there is dance language competence is resolved here by way of an examination of its theoretical efficacy; the dance language is demonstrably rule-guided rather than regular if the behavior can

be idealized along a dimension which may be plausibly described in knowledge/use, conceptual/perceptual terms, idealized in such fashion that there is a grammar accounting for the idealized behavior which is not only observationally and descriptively, but also "explanatorily" adequate.

"Explanatory" adequacy is the genetically coded counterpart to the explanatory adequacy demand made of grammar and grammatical theory with respect to human language (see Chomsky 1964), which of course is largely acquired, unlike the genetically coded dance language. Thus, while linguistic theory must characterize the boundaries of possible variation in human language so that the idiosyncratic grammatical features are learnable, predicting just what is a possible human language (giving the theory explanatory adequacy), dance linguistic theory must provide a basis for predicting just which aspects of the behavior are subject to "dialectal" variation, predicting just what is a possible dance language (giving the theory "explanatory" adequacy).

In the case of human linguistic theory, the residue after sifting the universals out of the grammar of a particular language is required to be learnable; see Levelt 1974b:39-41 for a lucid critique of current transformational theory on precisely these acquisition-based grounds.<sup>1</sup> In the case of dance linguistic theory, it is

not the residue after universal-titration which it is a challenge to constrain. The "dialect"-particular residue, unlike the universals, is easily viewed as a genetic unit, the behavioral correlate of which is the Y-shaped vs. sickle transitional dance option. In the event that dialectal variation in distance correlation is best viewed as linguistic (and not, as here, as a consequence of different energy drain to distance translations due to racial variation in bodily size), this dialectal parameter is also easily viewed as a genetic unit.<sup>2</sup> What it is a challenge to unify are the dance linguistic universals, those apparently widely diverse dependencies (diverse from the digital point of view of sec. 3, that is) observed to function for the "winged ganglion," the honey bee as a genetic unit interracially persistent in Apis.

This emphasis on "explanatory" adequacy reflects a concern with genetic packaging untinged by the virulent functionalism of the post Wenner years (see, for example, von Frisch 1968, Dawkins 1969, Lindauer 1971). It is essentially a static synchronic genetic investigation, with the emphasis on phylogenetic prediction rather than dynamic evolutionary reconstruction (see, for example, Esch 1967, Esch et al. 1965, von Frisch 1962).

The digital account of sec. 3 to a great extent captures the traditional view of the dance language, utilizing explicit linguistic terminology conducive to

evaluation along observational, descriptive, and "explanatory" adequacy dimensions. Consider, for example, the identification of the distance and direction correlations as linguistic, goal odor and quality (given by dance vivacity) correlations as paralinguistic. While this is not a distinction drawn in so many words in the experimental literature, it is clear that this distinction is a conceptual prerequisite to the formulation of the dance language controversy--and both distance and direction linguistic parameters have been shown to function in communication (Gould 1975b, Schricker 1974a).

The single major discrepancy between the view of sec. 3 and that traditionally assumed lies in the finitizing of the vocabulary, a move not so much traditionally resisted as left unexplored. On the one hand, there has been an enormous reluctance, on the whole without very much in the way of empirical evidence or supportive argumentation, to view the system as rule-guided; Bennett 1964 and Chomsky 1975b are, regrettably, quite representative.<sup>3</sup> On the other hand, there is considerable confusion as to the theoretical status of a mathematically digitalizing idealization of analog data. Souček and Carlson 1976 and Block and Fodor (1976(?)), for example, agree that such a digitalization is harmless since it may be drawn as finely as desired, i.e., since the  $n$  of  $\infty_1, \infty_2, \dots, \infty_n$  may be as large as you like, nothing is lost by the imposition

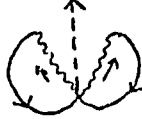
of an n-ary limit. This is quite simply false in a cognitive (competence) domain; the mathematical distinctions between finite and infinite, denumerable and nondenumerable carry conceptual if not perceptual weight. Thus if the dance language is rule-guided (as it is by assumption, that assumption still to be legitimized by an "explanatorily" adequate dance linguistic theory), the  $\aleph$  (and  $\omega$ ) cardinality issue is vital, the finite/denumerable/nondenumerable choice possibly difficult to motivate but crucially important. The linguistic analysis of the dance language of the honey bee thus subsumes the problem of determining the lexical cardinality conducive to the formalization of an "explanatorily" adequate theory.

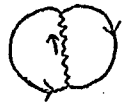
The digital analyses proffered in secs. 3-4.2 accept the traditional view of the dance (or superdance) message, viz., that the dance behavior uses distance and directional coding mechanisms to refer to the point site of a goal previously visited by the dancer. Apparently observationally adequate finite state accounts are given for a variety of plausible lexical patternings. With regard to the distance dimension,  $\omega$ , three possibilities are explored: (a) that there are a finite number of distance values,  $\omega_1, \omega_2, \dots, \omega_m$ ; (b) that there are a denumerable number of distance values because there can be no maximal  $\omega_m$ , no largest  $\omega$ -value at the cognitive level; and (c) that there are a denumerable number of  $\omega$ -values,

whether or not there is a cognitive maximum of  $\omega_m$ , because perceptual averaging of any  $i$ ,  $i \geq 2$ , of them must give an element of the cognitive set.  $\alpha$ , of course, is finite if digital. The last case, (c), is motivated by the position that it is superdance SD rather than dance D which is properly taken to be the subject matter of the syntactic component; the other two are independent of the D/SD issue and, in fact, the resulting grammars relatively easily accommodate either claim. Within the confines of the traditional view of the danced message, SD is, if not ignored entirely, most naturally taken to be a message repetition dance concatenation.

In sec. 4.1.2, the finite state approach is shown to be in all likelihood descriptively adequate, provided only that the output strings are simultaneously judiciously parsed via finite state generated labelled bracketings. Descriptive adequacy is a matter of capturing the dancer's structural "intuitions," these being accessible by way of an examination of the categorizing the operation parameters of the (noise-contributing) performance system suggest (see Fodor 1966).

Within the digital paradigm, residual misdirection (the  $\alpha$ -dependent angular error physiologically induced by the dark hive, vertical dance floor, by-gravity context) assumes a suspicious degree of importance with respect to the goal of descriptive adequacy. For, in the event that

in transitional dances  it is the direction-giving bisector which is residually misdirected, that bisector, as the input to a performance process, must have grammatical status. That grammatical status may be achieved syntactically (as by deriving transitional dances using trans-

formations spreading overlapping circuits  the appropriate  $\omega$ -dependent amount) or semantically (as by projection rules which average wagging runs), but it inevitably increases the resemblance between the dance language and human language. This follows even if finite state bracketing is used to capture the supposed bisector misdirection.

Sec. 2's analysis of the performance system, as a critical review of the literature, is compatible with, and to some extent dependent on, a competence account of the sort sketched in secs. 3-4.1.2, these being alternative formalizations of the traditional view of the dance language.

Arguments are presented in support of perceptual averaging at three stages of the message-coding communication process: (a) the perceptual integration of (lateral drift corrected?) flight percepts (contra von Frisch 1967a), a process which discards "outvoted" perceptual data (so that transport experiments stabilize with forager training), and generally prefers the outbound flight; (b) averaging of roughly six wagging runs in dance comprehension; with

SD understood in error-plagued message repetition terms; and (c) weighted averaging of the competing directional poles in light-dependent misdirection cases (contra von Frisch 1967a).

"Distance" coding is problematic at both the message (distance/energy drain/flight range percentage) and signal (duration of wagging/duration of sound production/length of wagging run etc.) levels. The need to distinguish the linguistic coding of "distance" from the definitive cue in communication which follows from the competence/performance distinction precludes reliance on results of projected model bee studies. It is similarly linguistically naive to suppose that production and comprehension of "distance" is accomplished by means of computational inverses operating with equal accuracy. It is argued that the determination of the linguistic coding of "distance" depends first of all on simplicity considerations and only secondarily on an energy drain accuracy comparison, restricted to physiologically interdependent candidates; the existence of physiologically independent measures surviving a simplicity scything entails multiple linguistically (grammatically) dignified "distance" codings. The distance/energy drain decision is related, in spirit at least, to the regular vs. rule-guided issue, with the rule-guided hypothesis currently under consideration urging the

traditional distance interpretation. This issue, unlike the distance/flight range percentage problem, does not affect the grammar as a formal system, but only the interpretation of that formal system. The distance/flight range percentage issue arises in secs. 3-4.1.2 as the question of whether the  $\omega$  "distance" range is finite; if  $\omega$  measures distance per se then at the competence level there is no maximum  $\omega_m$  and hence there are an infinite number of  $\omega$  values, while if  $\omega$  measures flight range percentage, there is a conceptual maximum at 100% of the flight range. Both lexical alternatives are explored.

The traditional view of the dance language message set significantly underdetermines the grammar, particularly at the lexical level, as has been outlined. However, although the analytical options are numerous, "explanatory" adequacy remains stubbornly elusive.

An "explanatorily" adequate theory must account for the fact there are different dance forms, identically ordered by each race with increasing distance, that the shape of the transitional dance series is alone racially variable, that wagging duration and angular deviation across SD correlate inversely, with the angular deviation maximizing at  $60^\circ$ , etc. The search area hypothesis of sec. 4.4 is just such an "explanatorily" adequate dance linguistic theory.

The move from a point source to a circular patch

of constant, genetically fixed radius at the message level, with SD coding the patch rather than the goal, accounts for all these observations. What's more, it does so by undermining the view of the dance language as a regular system, since it is not the food stimulus's location per se which is coded. Finally, the search area hypothesis is not compatible with any of the countable (i.e., finite or denumerable) "distance" lexicons which might be motivated. The conclusion: the dance language is a rule-guided animal communication system, syntactically different in kind rather than degree from human language.

The shift from a recursively enumerable set to one which by its very (nondenumerable) cardinality is not susceptible to grammatical analysis of the usual sort, forcing the definitional approach of, for example, (84) and (90), mandates an analysis of production and comprehension coding strategies. This is the major change in the performance model occasioned by the basic alteration in the competence account from a digital to an analog model.

Gould 1975a:171, presumably on the basis of simplicity considerations, suggests that production and comprehension are computational inverses operating with approximately equal accuracy in communication, detailing how this assumption would permit the identification of the definitive cue for "distance." Simplicity considerations combine with

the search area hypothesis to suggest production and comprehension strategies which are not only computational inverses but are moreover identical, actually entailing equal accuracy in communication.

In this regard, it is first shown how "explanatory" adequacy requires the assumption that directional but not distance variation in SD is linguistic--only this assumption predicts the relevant dance type to "distance" range correlation. The production and comprehension (Ballantine) strategies then are point-to-patch cognitive extension processes, coding a goal at  $P_0$  by means of the two directionally most distant patch boundary points  $P_L$  and  $P_R$ , recovering  $P_0$  from  $P_L$  and  $P_R$  by point to patch extensions about  $P_L$  and about  $P_R$ , which patches intersect at  $P_0$ . The overall effect of the search area hypothesis is to suggest two-dimensional geometric rather than algebraic (averaging) processing.

A good deal of what, on the digital accounts of secs. 3-4.2.1, was handled by brute force observationally unmotivated competence distinctions, surfaces here as perceptually motivated performance phenomena. The  $P_L/P_R$  Ballantine syntactic patch coding strategy accounts for the full-wagging transitional and tailwagging dances; the other dance types are not dance types at all, but reflect problems due to perceptual limitations hampering the execution of the strategy. The round dance, for

example, is prompted by goals such that the point-to-patch cognitive net includes the hive; the S-dance results from an angular deviation so small that it is beyond the sensory capabilities of the dancer to reliably produce. The transitional dance's  $60^\circ$  maximum in SD angular deviation results from an interaction of the Ballantine production and comprehension strategies. It is moreover due to a distance coding conflict, thus explaining why such dance forms maintain a direction correlation. The two lobes of the transitional dance reflect the two-strategy interaction nature of the problem, even as the one lobe shape of the round dance reflects its one-strategy source. The dance language is a good deal more iconic than could have been suspected within the traditional view of the message set.

The distance vs. flight range issue remains mysterious. Handling it within the grammar, whether by an  $\omega$  maximum in the digital framework or an R-clause (see (84)) in the analog framework, is to make the observation without explaining it. If it is in fact the case that the extreme 100% of the flight range figure is represented by the same degree of wagging duration for each race, then clearly this is something an "explanatorily" adequate theory must account for; a certain (racially invariant) real number multiple of the (racially dependent) patch radius  $c/2$  gives the flight range. Von Frisch 1967a: 303-304 casts some doubt on the percentage of flight

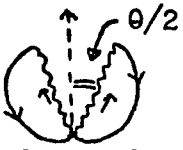
range "distance" interpretation, describing A. indica as having a flight range of 750m. (dancing for up to 700m.), A. florea as having one of 300-400m., yet A. indica apparently does direction-indicating dances for smaller distances than did A. florea: "In this reference to the direction of a goal that is only a few centimeters beyond the hive entrance A. indica stands alone." If this observation survives investigation in energy drain terms, then the flight range percentage "distance" interpretation is disconfirmed, and "distance" is literally distance at the rule-guided competence level of analysis.

The dance language controversy has served to channel dance linguistic research into the problem of recruitment, an area of investigation with communicative but little if any linguistic import. The search area hypothesis is most obviously vulnerable to experimental testing as to genetic dance language variation and the study of individual dancers and the dance correlations they maintain in production. With regard to the first of these, the points made here appear to be universally upheld in the literature. Production data, on the other hand, has unfortunately received very little attention in the literature.

Gould 1975b:192, in calculating the information content of the (digitalized) dance language reports that "the SD's [standard deviations] from my measurements and

those of others are  $16^\circ$  at 150m,  $5^\circ$  at 400m,  $4^\circ$  at 750m, and  $3^\circ$  beyond 1 km." These figures, as the  $\theta$  value for the given  $r_0$  value, accord well with the 20m round dance range of use displayed by his stock of Italian foragers (Gould 1976:213); the formulas of sec. 4.4 yield  $c/2$  values of 21.1m, 17.5m, 26.0m, 26.5m ( $r_0=1$  km), respectively, which average to 22.8m. There is some question, however, as to whether  $\theta$  is legitimately taken to be the angular SD [standard deviation].

The standard deviation is a measure of how close one can expect the value of the variable to be to the mean. With respect to the paradigm case of the search area

hypothesis, i.e. the transitional dance , the mean appears to refer to the bisector and thus, the standard deviation to  $\theta/2$ , not  $\theta$ . The obvious  $\theta$  values of  $32^\circ$  for 150m,  $10^\circ$  for 400 m, etc., however, produce outsized  $c/2$  (round dance range of use) values of 43m, 35m, etc.

This would act to disconfirm the search area hypothesis were it not for the puzzling report of Gould 1975c that the mean recruit error at 150m is  $11.9^\circ$  and at 400m is  $4.2^\circ$ . Then the sum-of-the-squares formula predicated on the computational inverses view of production and comprehension (which Gould endorses--see discussion, pp. 33-34--and which the matching strategies here outlined for production and comprehension account for) gives production

error ( $\theta/2$ ) values of  $8^{\circ}20'$  at 150m,  $3^{\circ}$  at 400m, i.e.  $\theta$  values of  $16^{\circ}40'$  and  $6^{\circ}$ , which certainly correlate better with the SD's of the previous paragraphs than with their doubles.

Ideally, what should be tabulated is the average angular deviation from the mean direction of the superdance for each wagging duration value.  $c/2$  is ideally given as a wagging duration figure for the individual dancer. The actual physical distance between the hive and the goal is irrelevant to this determination. Again, the theory is most vulnerable to SD [superdance] angular measurements taken for a single dancer, and, although it seems likely that the experimental work done thus far has produced the data relevant to an evaluation of the search area hypothesis, they simply do not figure in the published reports of that research.

Structural considerations strongly urge the search area grammatical hypothesis, primarily on grounds of "explanatory" adequacy. The conclusion drawn here, that the dance language is rule-guided but syntactically different in kind rather than degree from human language, awaits phylogenetic and experimental evaluation.

Now that the communicative status of the dance language has been experimentally established, it is its linguistic status which is in question. The search area hypothesis may well prove to be the answer to that question.

## Notes

5

## IMPLICATIONS, EXPERIMENTAL AND LINGUISTIC

1. Given a linguistic theory whose titration gives in every case learnable language particulars, the universals thereby isolated would require theoretical analysis, as in the dance linguistic case discussed below. At the present time, however, this is not by any means a major consideration.

2. A demonstration that these two parameters interact in some way would urge a reanalysis of these as not two, but somehow one linguistic unit; this seems unlikely, but cannot be dismissed out of hand. It is conceivable that sickle dancing is genetically tied to races of a certain body size, in which case considerations of "explanatory" adequacy would, other things being equal, favor a grammar somehow uniting these features linguistically.

3. Griffin 1976 is a notable exception to this practice.

## BIBLIOGRAPHY

- Akmajian, A., and F. Heny (1975). An Introduction to the Principles of Transformational Syntax. Cambridge, Mass.: The MIT Press.
- Allen, M.D. (1959a). The occurrence and possible significance of the 'shaking' of honeybee queens by the workers. Animal Behaviour 7 66-69.
- Allen, M.D. (1959b). The 'shaking' of worker honeybees by other workers. Animal Behaviour 7 233-240.
- Bach, E. (1974). Syntactic Theory. New York: Holt, Rinehart and Winston, Inc.
- Bellugi, U., and E. S. Klima (1975). Aspects of sign language and its structure. In Kavanagh and Cutting 1975, 171-203.
- Bennett, J. (1964). Rationality: Essay towards an Analysis. London: Routledge and Kegan Paul.
- Bisetzky, A. R. (1957). Die Tänze der Bienen nach einem Fussweg zum Futterplatz. Zeitschrift für vergleichende Physiologie 40 264-288. As cited in von Frisch 1967a.
- Block, N., and J. A. Fodor (1976?). Cognitivism and the analog/digital distinction. Unpublished paper, MIT Department of Linguistics and Philosophy.
- Boch, R. (1957). Rassenmassige Unterschiede bei den Tänzen der Honigbiene (Apis mellifera L.). Zeitschrift für vergleichende Physiologie 40 289-320. As cited in von Frisch 1967a.
- Bossert, W. H., and E. O. Wilson (1963). Analysis of olfactory communication among animals. Journal of Theoretical Biology 5 443-469.
- Bronowski, J. (1967). Human and animal languages. In To Honor Roman Jakobson, v.I., The Hague: Mouton, 374-395.

- Bronowski, J., and U. Bellugi (1970). Language, name, and concept. *Science* 168 669-673.
- Chomsky, N. (1959). On certain formal properties of grammars. *Information and Control* 2 137-167.
- Chomsky, N. (1963). Formal properties of grammars. In Luce et al. 1963, 323-418.
- Chomsky, N. (1964). Current Issues in Linguistic Theory. The Hague: Mouton.
- Chomsky, N. (1965). Aspects of the Theory of Syntax. Cambridge, Mass.: The MIT Press.
- Chomsky, N. (1975a). Reflections on Language. New York: Random House.
- Chomsky, N. (1975b). Knowledge of language. In Gunderson 1975, 299-320.
- Chomsky, N., and G.A. Miller (1963). Introduction to the formal analysis of natural languages. In Luce et al. 1963, 269-322.
- Dawkins, R. (1969). Bees are easily distracted. *Science* 165 751.
- Dethier, V.G. (1957). Communication by insects, physiology of dancing. *Science* 125 331-336.
- Esch, H. (1961) Ein neuer Bewegungstyp im Schwänzeltanz der Bienen. *Naturwissenschaften* 48 140-141. As cited in von Frisch 1967a.
- Esch, H. (1963). Über die Auswirkung der Futterplatzqualität auf die Schallerzeugung im Werbetanz der Honigbiene. *Verhandlungen der Deutschen zoologischen Gesellschaft in Wien 1962 (Leipzig, 1963)* 302-309. As cited in von Frisch 1967a.
- Esch, H. (1964). Beiträge zum Problem der Entfernungsweisung in den Schwänzeltänzen der Honigbienen. *Zeitschrift für vergleichende Physiologie* 48 534-546. As cited in von Frisch 1967a.
- Esch, H. (1967). The evolution of bee language. *Scientific American* 216 (4) 96-103.
- Esch, H., and J. A. Bastian (1970). How do newly recruited honey bees approach a food site? *Zeitschrift für vergleichende Physiologie* 68 (2) 175-181.

- Esch, H., I. Esch, and W. E. Kerr (1965). Sound: an element common to communication of stingless bees and to dances of honey bees. *Science* 149 320-321.
- Ferguson, A. (1975). Evolution von Frisch, and teleology. *The American Naturalist* 109 (1967), 369-370.
- Finke, I. (1958). Zeitgedächtnis und Sonnenorientierung der Bienen. Lehramtsarbeit Naturw. Fak. Univ. München. As cited in von Frisch 1967a.
- Fodor, J. A. (1966). Discussion of Kalmus 1966. In Smith and Miller 1966, 287-294.
- Fouts, R.S., and J.B. Couch (1976). Cultural evolution of learned language in chimpanzees. In Hahn and Simmel 1976, 141-161.
- Frisch, K. v. (1962). Dialects in the language of bees. *Scientific American* 207 78-87.
- Frisch, K. v. (1967a). The Dance Language and Orientation of Bees. Cambridge, Mass.: Harvard University Press.
- Frisch, K. v. (1967b). Honey bees: do they use the direction and distance information provided by their dancers? *Science* 158 1072-1076.
- Frisch, K. v. (1968). Do bees really not understand their own language? *Animal Behaviour* 16 531-533.
- Frisch, K. v., and R. Jander (1957). Über den Schwänzeltanz der Bienen. *Zeitschrift für vergleichende Physiologie* 40 239-263.
- Frisch, K. v., and M. Lindauer (1954). Himmel und Erde in Konkurrenz bei der Orientierung der Bienen. *Naturwissenschaften* 41 245-253. As cited in Lindauer 1967.
- Frisch, K. v., M. Lindauer, and F. Schmeidler (1960). Wie erkennt die Biene den Sonnenstand bei geschlossener Wolkendecke? *Naturwissenschaftliche Rundschau* 169-172. As cited in von Frisch 1967a.
- Gahl, R. A. (1975). The shaking dance of honey bee workers; evidence for age discrimination. *Animal Behaviour* 23 230-232.
- Gardner, B. F., and R.A. Gardner (1971). Two-way communication with an infant chimpanzee. In Schrier and Stollnitz 1971, 117-185.

- Gary, N.E., and P.C. Witherell (1971). A method for training honey bees to forage at feeding stations. Annals of the Entomological Society of America 64 (2) 448-449.
- Goncalves, L.S. (1969). A study of the orientation information given by one trained bee by dancing. Journal of Apicultural Research 8 113-132.
- Gould, J. L. (1974). Honey bee communication. Nature 252 300-301.
- Gould, J. L. (1975a) Honey bee recruitment: the dance-language controversy. Science 189 685-693.
- Gould, J. L. (1975b). Honey bee communication: the dance-language controversy. Thesis, Rockefeller University.
- Gould, J. L. (1975c). Communication of distance information by honey bees. The Journal of Comparative Physiology 104 161-173.
- Gould, J. L. (1976). The dance language controversy. The Quarterly Review of Biology 51 (2), 211-244.
- Gould, J. L., M. Henerey, and M.C. MacLeod (1970). Communication of direction by the honey bee. Science 169 544-554.
- Griffin, D.R. (1976). The Question of Animal Awareness: Evolutionary Continuity of Mental Experience. New York: The Rockefeller University Press.
- Gunderson, K. (ed.) (1975). Language, Mind and Knowledge. Minneapolis: University of Minnesota Press.
- Hahn, M.E., and E. C. Simmel (eds.) (1976). Communicative Behavior and Evolution. New York: Academic Press.
- Haldane, J.B. S., and H. Spurway (1954). A statistical analysis of communication in Apis mellifera and a comparison with communication in other animals. Insectes Sociaux 1 247-283.
- Harnad, S. R., H.D. Steklis, and J. Lancaster (eds.) (1976). Origins and Evolution of Language and Speech. New York: New York Academy of Sciences Annal 280.

- Fealy, A. F. (1973). Can chimpanzees learn a phonemic language? *Journal of Psycholinguistic Research* 2 (2) 167-170.
- Heran, H. (1955). Versuche über die Windkompensation der Bienen. *Naturwissenschaften* 42 132-133. As cited in von Frisch 1967a.
- Heran, H. (1956). Ein Beitrag zur Frage nach der Wahrnehmungsgrundlage der Entfernungsweisung der Bienen. *Zeitschrift für vergleichende Physiologie* 38 168-218. As cited in von Frisch 1967a.
- Heran, H., and M. Lindauer (1973). Windkompensation und Seitwindkorrektur der Bienen beim Flug über Wasser. *Zeitschrift für vergleichende Physiologie* 47 39-55. As cited in von Frisch 1967a.
- Herman, G.T., and G. Rozenberg (1975). Developmental Systems and Languages. Amsterdam: North Holland.
- Hockett, C. F. (1960). Logical considerations in the study of animal communication. In Lanyon and Tavolga 1960, 392-427.
- Hockett, C.F., and S. A. Altmann (1968). A note on design features. In Sebeok 1968, 61-72.
- Janda, A. (1977). Concerning the syntactic analysis of the dance language of the Carniolan honeybee. *CUNYForum Papers in Linguistics* 2 132-158.
- Jander, R. (1957). Die optische Richtungsorientierung der roten Waldameise (Formica rufa L.). *Zeitschrift für vergleichende Physiologie* 40 162-238. As cited in von Frisch 1967a.
- Johnson, D. L. (1967a). Communication among honey bees with field experience. *Animal Behaviour* 15 487-492.
- Johnson, D. L. (1967b). Honey bees: do they use the direction information contained in their dance maneuver? *Science* 155 844-847.
- Johnson, D. L., and A. M. Wenner (1966). A relationship between conditioning and communication in honey bees. *Animal Behaviour* 14 261-265.
- Johnson, D. L., and A. M. Wenner (1970). Recruitment efficiency in honeybees: studies on the role of olfaction. *Journal of Apicultural Research* 9 (1) 13-18.

- Kalmus, H. (1956). Sun navigation of Apis mellifica L. in the southern hemisphere. Journal of Experimental Biology 33 554-565.
- Kalmus, H. (1966). Ontogenetic, genetical, and phylogenetic parallels between animal communication and prelinguistic child behaviour. In Smith and Miller 1966, 273-285.
- Katz, J. J. (1972). Semantic Theory. New York: Harper and Row.
- Katz, J. J. (1976). A hypothesis about the uniqueness of natural language. In Harnad, Steklis and Lancaster 1976, 33-41.
- Katz, J. J. (1977). The real status of semantic representations. Linguistic Inquiry 8 (3) 559-584.
- Katz, J. J., and D. T. Langendoen (1976). Pragmatics and presupposition. Language 52 (1) 1-17.
- Kavanagh, J. F., and J. E. Cutting (eds.) (1975). The Role of Speech in Language. Cambridge, Mass.: The MIT Press.
- Konishi, M. (1970). Evolution of design features in the coding of species-specificity. American Zoologist 10 67-72.
- Krames, L., P. Pliner, and T. Alloway (eds.). (1974). Nonverbal Communications. New York: Plenum Press.
- Kroeber, A. L. (1952). Sign and symbol in bee communications. Proceedings of the National Academy of Sciences 38 (9) 753-757.
- Langendoen, D. T. (1976). On the adequacy of type-3 and type-2 grammars for human languages. CUNYForum Papers in Linguistics 1 1-12.
- Lanyon, W. E., and W. N. Tavolga (eds.). (1960). Animal Sounds and Communication. Washington, D.C.: Intelligence Printing Co.
- Levelt, W. J. M. (1974a). Formal Grammars in Linguistics and Psycholinguistics (vol. I): An Introduction to the Theory of Formal Languages and Automata. The Hague: Mouton.

- Levelt, W. J. M. (1974b). Formal Grammars in Linguistics and Psycholinguistics (vol. II): Applications in Linguistic Theory. The Hague: Mouton.
- Limber, J. (1976). Unravelling competence, performance and pragmatics in the speech of young children. *Journal of Child Language* 3 309-318.
- Limber, J. (1977). Language in child and chimp? *American Psychologist* 32 (4) 280-295.
- Lindauer, M. (1957). Zur Biologie der stachellosen Bienen, ihre Abwehrmethoden. Bericht 8. Wanderversammlung Deutscher Entomologen (Berlin) 71-78. As cited in von Frisch 1967a.
- Lindauer, M. (1959). Angeborene und erlernte Komponenten in der Sonnenorientierung der Bienen. *Zeitschrift für vergleichende Physiologie* 42 43-62. As cited in von Frisch 1967a.
- Lindauer, M. (1963). Kompassorientierung. *Ergebnisse der Biologie* 26 158-181. As cited in Gould 1976.
- Lindauer, M. (1967). Communication among Social Bees. New York: Atheneum Press.
- Lindauer, M. (1971). The functional significance of the honeybee waggle dance. *The American Naturalist* 105 (942) 89-96.
- Lindauer, M. (1973). Das Magnetfeld der Erde als Orientierungshilfe für die Bienen. *Imkerfreund* 28 3-7. As cited in Gould 1976.
- Lindauer, M. (1974). Die Orientierung der Bienen-neue Erkenntnisseneue Rätsel. *Biene* 110 134-138. As cited in Gould 1976.
- Lindauer, M., and H. Martin (1968). Die Schwereorientierung der Bienen unter dem Einfluss des Erdmagnetfeldes. *Zeitschrift für vergleichende Physiologie* 60 219-243. As cited in Gould 1976.
- Lindauer, M., and B. Schrickler (1963). Über die Funktion der Ocellen bei den Dämmerungsflügen der Honigbiene. *Biologisches Zentralblatt* 82 721-725. As cited by Gould 1975b.
- Luce, R.D., R. R. Bush, and E. Galanter (eds.) (1963). Handbook of Mathematical Psychology, v. 2. New York: John Wiley.

- McCulloch, W. S., and W. Pitts (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5 115-133.
- Markl, H. (1966). *Zeitschrift für vergleichende Physiologie* 53 328-352.
- Marler, P. (1970). Bird song and speech development: could there be parallels? *American Scientist* 58 669-673.
- Martin, H., and M. Lindauer (1973). Orientierung im Erdmagnetfeld. *Fortschritte der Zoologie* 21 211-228. As cited in Gould 1976.
- Mautz, D. (1971). Der kommunikationseffekt der Schwänzel-tanze bei *Apis mellifica carnica* (Pollm.). *Zeitschrift für vergleichende Physiologie* 72 (2) 197-220. As cited in Gould 1975b.
- Mistler-Lachman, J. L., and R. Lachman (1974). Language in man, monkeys, and machines. *Science* 185 871-872.
- New, D. A. T. (1961). Effects of small zenith distances of the sun on the communication of honey bees. *Journal of Insect Physiology* 6 196-208.
- New, D. A. T., and J. K. New (1962). The dances of honey bees at small zenith distances of the sun. *Journal of Experimental Biology* 39 271-291.
- Nottebohm, F. (1968). The 'critical period' for song learning. *Ibis* 111 386-387.
- Nottebohm, F. (1971). Neural lateralization of vocal control in a passerine bird. I. Song. *The Journal of Experimental Zoology* 177 229-262.
- Nottebohm, F. (1972). Neural lateralization of vocal control in a passerine bird. II. Subsong, calls, and a theory of vocal learning. *The Journal of Experimental Zoology* 179 35-50.
- Nottebohm, F. (1975). A zoologist's view of some language phenomena with particular emphasis on vocal learning. In *Foundations of Language Development v.1*. New York: Academic Press, 61-101.
- Nottebohm, F., and M. E. Nottebohm (1976). Left hypoglossal dominance in the control of canary and white-crowned sparrow song. *Journal of Comparative Physiology A* 108 171-192.

- Otto, F. (1959). Die Bedeutung des Rückfluges für die Richtungs- und Entfernungsangabe der Bienen. Zeitschrift für vergleichende Physiologie 42 303-333. As cited in von Frisch 1967a.
- Pankova, S. V. (1972). About signal dances of honey bees. In XXIIIrd International Apicultural Congress. Bucharest, Romania: Apimondia Publishing House, 405-418.
- Peters, P. S., and R. W. Ritchie (1973). On the generative power of transformational grammars. Information Sciences 6 49-83.
- Postal, P. (1963). Some syntactic rules in Mohawk. Thesis, Yale University.
- Quine, W. V. O. (1960). Word and Object. Cambridge, Mass.: MIT Press, esp. 26-79.
- Rumbaugh, D., and S. Savage-Rumbaugh (1978). Chimpanzee language research: Status and potential. Behavior Research Methods and Instrumentation 10 (2) 119-131.
- Schlesinger, I. M. (1968). Sentence Structure and the Reading Process. The Hague: Mouton.
- Scholze, E., H. Pichler, and H. Heran (1964). Zur Entfernungsschätzung der Bienen nach dem Kraftaufwand. Naturwissenschaften 51 69-70. As cited in von Frisch 1967a.
- Schricker, B. (1965). Die Orientierung der Honigbiene in der Dämmerung. Zeitschrift für vergleichende Physiologie 49 420-458. As cited by Gould 1975b.
- Schricker, B. (1947a). Der Einfluss subletaler Dosen von Parathion (E605) auf die Entfernungsweisung bei der Honigbiene. Apidologie 5 (2) 149-175.
- Schricker, B. (1974b). Der Einfluss subletaler Dosen von Parathion (E605) auf das Zeitgedächtnis der Honigbiene. Apidologie 5 (4) 385-398.
- Schricker, B., and W. P. Stephen (1970). The effect of sublethal doses of parathion on honeybee behavior. I. Oral administration and the communication dance. Journal of Apicultural Research 9 (3) 141-153.
- Schrier, A. M., and F. Stollnitz (eds.) (1971). Behavior of Nonhuman Primates, v. 4. New York: Academic Press.

- Sebeok, T. A. (ed.) (1968). Animal Communication: Techniques of Study and Results of Research. Bloomington: Indiana University Press.
- Silva De Moraes, R. L. M., and C. da Cruz Landem (1972). Estudo comparativo de orgaos sensoriais em abelhas com diferentes tipos de comunicacao. *Revista Brasileira de Biologia* 32 (2) 185-196.
- Smith, F., and G. A. Miller (eds.) (1966). The Genesis of Language: a Psycholinguistic Approach. Cambridge, Mass.: The MIT Press.
- Smith, W. J. (1969). Messages of vertebrate communication. *Science* 165 145-150.
- Souček, B., and A. D. Carlson (1976). Computers in Neurobiology and Behavior. Part II: Computer Models of Neural Activities and of Animal Communications. New York: John Wiley.
- Steche, W. (1954). Gibt es "Dialekte" der Bienensprache? Diss. Naturw. Fak. Univ. München. As cited in von Frisch 1967a.
- Stephen, W. P., and B. Schricker (1970). The effect of sublethal doses of parathion. II. Site of parathion activity, and signal integration. *Journal of Apicultural Research* 9 (3) 155-164.
- Stokol, W., D. Casterline, and C. Croneberg (1965). A Dictionary of American Sign Language on Linguistic Principles. Washington, D.C.: Gallaudet College Press.
- Thorpe, W. H. (1974). Animal Nature and Human Nature. Garden City, New York: Anchor Press.
- Valian, V. (1976). The relationship between competence and performance: a theoretical review. *CUNYForum Papers in Linguistics* 1 64-101.
- Wells, P. H., and A. M. Wenner (1971). The influence of food scent on behavior of foraging honey bees. *Physiological Zoology* 44 (4) 191-209.
- Wenner, A. M. (1961). A method of training bees to visit a feeding station. *Bee World* 42 8-11.

- Wenner, A. M. (1967). Honey bees: do they use the distance information contained in their dance maneuver? *Science* 155 847-849.
- Wenner, A. M. (1974). Information transfer in honeybees: a population approach. In Krames et al. 1974, 133-170.
- Wenner, A. M., P. H. Wells, and D. L. Johnson (1969). Honey bee recruitment to food sources: olfaction or language? *Science* 164 84-86.
- Wenner, A. M., P. H. Wells, and F. J. Rohlf (1967). Analysis of the waggle dance and recruitment in honey bees. *Physiological Zoology* 40 317-344.