Western Journal of Communication, 67(1) (Winter 2003) 56–73

Human Affection Exchange: IV. Vocalic Predictors of Perceived Affection in Initial Interactions

Kory Floyd and George B. Ray
Arizona State University and Cleveland State University

This investigation examines the relationship between speakers' vocal characteristics and receivers' and observers' perceptions of speakers' affectionate intentions toward receivers in initial interactions. Hypotheses drawn from affection exchange theory were tested in an experimental procedure involving 48 triads of previously unacquainted young adults. Results revealed that (1) speakers' fundamental frequency was linearly related to observers' perceptions of speakers' affection level; (2) speakers' variation in fundamental frequency was linearly related to receivers' and observers' perceptions of speakers' affection level; (3) speakers' vocal intensity was unrelated to perceptions of their affection level; and, (4) speakers' fundamental frequency interacted with speakers' sex to influence receivers' perceptions of speakers' affection level.

THE COMMUNICATION OF affection is critical to the development and maintenance of personal relationships. Affectionate expressions often serve as critical incidents by which relational development is ascertained, while their absence may be taken as evidence of disinterest or relational deterioration (see Owen, 1987). While commonly associated with romantic relationships, affectionate communication is also common in friendships and family relationships (Floyd & Morman, 1997; Salt, 1991), and is even important in initial interactions due to its ability to contribute to relational development (see Floyd & Burgoon, 1999).

Crucial to the ability of affectionate communication to contribute to human relationships is the ability of humans to decode affectionate

KORY FLOYD (Ph.D., University of Arizona, 1998) is associate professor of human communication at Arizona State University and GEORGE RAY (Ph.D., University of Washington, 1983) is associate professor of communication at Cleveland State University. The authors are grateful for the assistance of Diane Doskoch, Matt Hurd, Claudia Owens, Dan Smedley, Bernie Strong, Keri Vandeusen, and Ryan Weyls. Address correspondence to the first author at College of Public Programs, Hugh Downs School of Human Communication, Arizona State University, PO Box 871205, Tempe AZ 85287-1205. (480) 965-3568. Email: kory@asu.edu
Editor's note: This article was initially reviewed by the previous editor, David Henry, and reviewers he selected.

behaviors accurately. That is, without being able to discriminate between behaviors that connote affection and behaviors that connote hostility, aggression, or simply disinterest, humans would not only have difficulty forming and maintaining significant relationships but they would deprive themselves of the physical, emotional, and financial resources that accompany such relationships. As part of an ongoing series of studies examining the nonverbal encoding and decoding of affection (e.g., Floyd, 1999a; Floyd & Morman, 1997, 1998, 2000), we focus our attention here on the use of the human voice to communicate affectionate intentions.

Of particular issue in these studies, and indeed in many studies of nonverbal cues, is the extent to which perceptual outcomes of nonverbal behaviors are resident in the behaviors themselves or are strongly influenced by the characteristics of a given social group or context. Theories and perspectives used in the study of nonverbal communication take different positions on this issue. On one end of the theoretic spectrum, Heider's (1958) principle of meaning embeddedness posits that meanings are not inherent in behaviors but, rather, that behaviors derive their meanings from the characteristics of a given social context. This approach presupposes a good deal of variability in the meanings ascribed to the same behavior when enacted in varying social situations. A more moderate approach is provided by perspectives such as Burgoon and Newton's (1991) social meaning model, which predicts that interpretations of particular nonverbal behaviors show relative consistency across social contexts, but *within* social groups. That is, specific behaviors (e.g., an emblematic gesture) are ascribed meaning by a social or cultural group and should be decoded fairly consistently by the members of that group, regardless of the particulars of the context in which they occur (see, e.g., Burgoon & Le Poire, 1999; Burgoon & Walther, 1990). However, the social meaning model would not necessarily posit such consistency *across* cultural or social groups.

At the other end of the spectrum are perspectives positing that, even across cultural or social groups, particular behaviors will be interpreted similarly because they are part of a broader "human behavioral vocabulary." Most visible among these perspectives are those that are grounded in Darwin's (1959) theory of evolution by means of natural selection. Darwin's theory posits that heritable characteristics—whether physical, behavioral, or psychological—will be passed from generation to generation at greater frequencies if they contribute to survival and procreation than if they do not. From this point of view, one may expect to find consistency between social and cultural groups in the meanings ascribed to particular behaviors, if a functional link between those behaviors and specific survival or procreation needs can be identified. For instance, the scowl is associated with disgust because that particular facial expression constricts nasal passages, restricting the intake of potentially noxious fumes or gases, thereby contributing to one's survival in the face of that potential danger. Darwinians would thus expect that the scowl would be decoded as communicating disgust regardless of one's cultural group or social upbringing (see Darwin, 1872/1965).

The present study is situated at the latter end of this theoretic spectrum. Drawing from affection exchange theory, we posit that particular vocalic characteristics are associated with perceived affection because of functions that such characteristics may have served toward the survival and procreation of humans' evolutionary ancestors. Drawing from this theoretic tradition does not cause one to disregard the effects of culture or socialization, but rather to look at aspects of behavior that may transcend such influences. A more detailed delineation of affection exchange theory is provided subsequently.

Affection Exchange Theory

Affection exchange theory (AET: Floyd, 2001; Floyd & Morman, 2001, in press) conceives of affectionate communication as an adaptive behavior that contributes to humans' superordinate motivations for viability (survival) and fertility (procreation). Assumed in the theory is the Darwinian principle of selective fitness, whereby those organisms best adapted to the demands of their environments are the most likely to survive and procreate. AET makes explicit links between the communication of affection and human viability and fertility. Specifically, affectionate communication is posted to increase survival chances because it contributes to the development and maintenance of human pair bonds, exposing one to their associated resources such as food and shelter (Postulate 1). Moreover, affectionate communication is posited to increase individuals' reproductive opportunities by signaling to potential sexual partners that one would be a fit parent (Postulate 2). The third postulate predicts that individuals' long-term fertility motivations are further served when they communicate affection to their biological children, because the benefits associated with receiving affection make the children more suitable as mates, increasing the chances that the children will themselves reproduce and will pass on their parents' genes to yet a new generation.

In these three ways, AET predicts that affectionate communication contributes directly to humans' superordinate goals. Because affection exchange is an adaptive behavior, then according to AET, it is also governed by the very motivations it serves. This is an important point because it recognizes that the adaptive function of affection is dependent on humans' evolved abilities to decode expressions of affection accurately and to discriminate between affectionate and nonaffectionate communication behaviors. Among humans' primitive ancestors, those with the sharpest abilities to tell friend from foe and to recognize the sights and sounds of potential danger were certainly those best able to survive. AET provides that modern humans similarly benefit to the extent that they possess these discriminative abilities, among which is the ability to accurately decode vocal sounds to determine whether the source has affectionate or aggressive intentions toward the receiver.

A compelling example of the use of specific vocal characteristics when encoding affectionate messages concerns the use of "babytalk" or "parentese." Several studies have documented that in affectionate interactions with romantic partners, humans have a tendency to adopt a vocalic pattern that mimics the way most humans talk to babies (Bombar & Littig, 1996; Ferguson, 1977; Garnica, 1977; Zebrowitz, Brownlow, & Olson, 1992). Babytalk has been observed in multiple cultures in North America, Asia, Europe, and Africa and is practiced by both men and women (whether or not they are parents) and by children (see, e.g., Ferguson, 1964; Fernald & Simon, 1984; Shute & Wheldall, 1989; Toda, Fogel, & Kawai, 1990). Apart from its linguistic features, which include the use of idioms and pet names, simplified sentence structures, and word repetitions, babytalk is also characterized by increased pitch, increased pitch variance, and decreased amplitude or a "softening" of the voice (Fernald & Simon, 1984; Zebrowitz et al., 1992). Incidentally, several studies suggest that vocalic characteristics of babytalk, rather than its linguistic features, have the strongest effect in eliciting positive affect from the hearer (Fernald, 1989, 1993; Werker & McLeod, 1989).

Are these vocalic characteristics—increased pitch, increased pitch variance, decreased amplitude—associated with loving or affectionate messages only in the context of interacting with someone with whom a loving relationship already exists, such as one's child or one's romantic partner? Or is it the case that these vocal characteristics carry affectionate connotations that humans are tuned in to acknowledge, whether in an established relationship or not? AET argues the latter: that there are specific, adaptive reasons why humans use these particular vocal patterns when speaking with loved ones and that these patterns are recognized as connoting affection whether they are occurring in established relationships or not. Below, we apply the logic behind AET to each of these vocal characteristics, review relevant empirical findings, and advance hypotheses for the current investigation.

Fundamental Frequency

Fundamental frequency (F_0) is the acoustic measure of the number of sound wave vibrations produced per second by a source of sound, such as a musical instrument or a set of vocal cords. Humans perceive F_0 as *pitch* and judge it along a continuum of high to low. When a source produces a high number of sound wave vibrations per second, the sound is perceived by the human ear to be a high-pitched sound;

conversely, low-pitched sounds are produced when a source emits fewer sound waves per second.

The F_0 of any sound, including that of the human voice, is dependent on the length of the apparatus producing the sound waves. When a piano cord is struck, for instance, the length of the cord determines the pitch of the resulting sound, with shorter cords producing more sound wave vibrations per second and, thus, a higher-pitched sound than that produced by longer cords. The F_0 of human voices is similarly dependent on the length of the vocal folds; those with shorter vocal folds (such as children) have voices with a higher modal pitch than those with longer vocal folds (such as adults).

With respect to the voice, the general pattern is that the larger the organism is, the longer its vocal folds and the lower its vocal F_0 are. This point is important because it provides that the modal pitch of an organism's voice is an indicator of the organism's physical size (and, by extension, of that organism's potential for physical threat). Physically small organisms, such as human babies, housecats, or small birds, produce vocal sounds that are high in F_0 and that, according to the evolutionary logic behind AET, should be perceived by larger organisms as unthreatening. Contrariwise, physically large organisms, such as large dogs, bears, or lions, produce vocal sounds that are low in F_0 and should be perceived by smaller organisms as signaling at least the potential for threat. Among primitive human ancestors, the ability to draw inferences from vocal pitch to physical size and potential threat would certainly have been adaptive, in that it would have allowed those with this ability to avoid potential harm from animals whose physical size would make them a threat.

This reasoning leads to the several conclusions, one of which is that higher-pitched voices should be perceived by humans as signaling intentions that are friendly rather than aggressive. Certainly, humans encode vocal messages of friendliness, nurturance, and affection using high-pitched voices. As mentioned, a prime example of this pattern is the use of babytalk, which is characterized by higher-than-normal vocal pitch (Fernald & Simon, 1984) and which is used when speaking to babies, pets, romantic partners, or others to whom affection is expressed. It follows logically that humans would also decode higher-pitched voices as expressing more affection than lower-pitched voices. Indeed, there is research to indicate that lower-pitched voices are perceived as expressing messages of dominance and aggression more so than their higher-pitched counterparts (Buller & Burgoon, 1986; Ohala, 1982).

Based on this reasoning, we predict that F_0 is positively related to perceived affection. Stated as a formal hypothesis:

H1: Speakers' fundamental frequency (F_0) is directly related to receivers' (a) and observers' (b) perceptions of their affection level toward receivers.

In addition to examining central tendencies on F_0, we also examined the effects of its variation. Research indicates that pitch variation is perceived by hearers as a pleasant vocal characteristic, whereas the lack of pitch variation (leading to a monotonic speaking style) is perceived as unpleasant (Buller & Burgon, 1986). Because humans must make judgments about how much they like, feel safe with, and trust others when deciding whether to form or forego relationships, characteristics that make one appear more pleasant, friendly, and affiliative should be positively associated with perceptions of liking and affection. It follows, therefore, that a vocal characteristic like pitch variation should be decoded as expressing messages of affection. This reasoning leads to our second hypothesis:

H2: Speakers' variance in fundamental frequency is directly related to receivers' (a) and observers' (b) perceptions of their affection level toward receivers.

Acoustic Intensity

The third vocalic property we investigated was acoustic intensity, a measure of the acoustic energy emitted vocally. Intensity is amplitude that is measured in raw units of acoustic energy rather than in decibels. As with fundamental frequency, the adaptive advantage of sensitivity to acoustic intensity concerns primitive human ancestors' abilities to recognize potential danger in the form of predators. The amount of acoustic intensity that an apparatus can produce is directly related to its size—for instance, playing a tuba produces more acoustic energy than playing a flute, even when the same amount of air is blown into each. Similarly, large organisms (with larger vocal folds) produce more sound than do small organisms (with smaller vocal folds). Because large animals are generally better able than small animals to threaten humans, then those humans with the evolved ability to recognize that loud vocal sounds are associated with greater danger than quiet vocal sounds were in the best position to avoid danger. Working from this reasoning (although not from AET), Tusing and Dillard (2000) predicted that louder human voices are associated with greater perceptions of interpersonal dominance than are quieter human voices. This hypothesis was supported in their investigation, as it had been in previous studies (e.g., Aronovitch, 1976; Buller & Burgoon, 1986; Harrigan, Gramata, Lucic, & Margolis, 1989). Kimble, Forte, and Yoshikawa (1981) further reported that a louder voice is associated with negative affect (see also Ohala, 1984), other things being equal.

The obverse of this logical chain, of course, is that quieter vocal sounds ought to be associated not with perceptions of dominance and danger but with perceptions of nurturance, safety, and affection. Humans have a tendency to decrease their vocal intensity when encoding affectionate verbal messages, as evidenced by research on babytalk (e.g., Fernald & Simon, 1984; Zebrowitz et al., 1992). Ray (1986) also reported that lower levels of loudness were associated with increased social attractiveness. It follows that a tendency to *decode* decreased vocal intensity as a sign of affection would be adaptive, and therefore humans should associate lower levels of vocal intensity with greater perceived affection. This leads to our final hypothesis:

H3: Speakers' acoustic intensity is inversely related to receivers' (a) and observers' (b) perceptions of their affection level toward receivers.

Method

Participants

Participants were 144 adults (72 male, 72 female) recruited from communication courses at a medium-sized university in the Midwestern United States. Participants ranged in age from 18 to 51, with a mean age of 23.14 years ($SD = 6.07$). Most (68%) were Caucasian, while 21% were African–American, 8% were Native American, 4% were Hispanic, 1% were Asian, and 14.5% were of other ethnic origins. Most (90%) had never been married, while 5% were married and 5% were divorced at the time of the study. Participation was voluntary and earned extra course credit.

Procedure

Participants, who were recruited for "a study of problem-solving techniques," signed up in same-sex triads for hour-long experimental sessions. We used previously unacquainted triads, rather than triads representing established relationships, to reduce the extent to which relational history might influence participants' or observers' perceptions of the confederates. In an established friendship or family relationship, for instance, one might have knowledge about the idiosyncratic ways in which another expresses affection, or the lack of it, and we wanted to free the triads in the current study from the influences of such relational knowledge. Upon reporting to the laboratory facility, participants were randomly assigned to the roles of confederate (C), naïve participant (P), and observer (O), and were told that C and P would be engaging in two videotaped problem-solving activities that O would observe. O was then ushered to the observation corridor, the site from which the experimental interactions were videotaped. C and P were situated in the interaction area of the laboratory, a small room with a coffee table, two swivel chairs, and a remotely operated video camera mounted on an upper corner of the wall. C and P were given an envelope containing two index cards and were told that each card described a problematic situation or dilemma. They were instructed to begin their activity by reading the first card aloud, discussing alternatives for addressing the problem described, and trying to reach consensus on the best way to solve the problem. They were also told that the activity would be timed and that the researcher would indicate by a knock on the door when they were to stop discussing the first problem and begin discussing the second. The problems, adapted from Hale and Burgoon (1984), dealt with (1) the theft of a friend's valuables by a sibling; (2) one's Catholic friend who is contemplating an abortion; (3) the infidelity of a best friend's fiancée; and (4) the impending visit of a cohabiting couple's unsuspecting parents. These situations were selected because of their demonstrated utility in generating conversation (see Floyd, 1999b; White & Burgoon, 1997). The topics were presented in a cyclical, counterbalanced order within conditions.

C and P were left alone to discuss the first problem situation, which was observed by O on the television monitor in the observation corridor. After two minutes, the researcher knocked on the door, indicating to C and P that they should move on to the second topic. After two more minutes, the researcher re-entered the room, stopped the activity, and told C and P that they were to complete some measures regarding "how well you think this interaction went." Under the guise that it would prevent them from seeing each other's answers, C and P were separated to complete these measures and C was ushered back to the reception area. P remained in the interaction area and completed postmeasures that included manipulation checks. O completed the same postmeasures.

C first completed manipulation check measures, was then asked to be the confederate, and was administered the behavior manipulation. C and P were then reunited in the interaction area, were given a new pair of problems to discuss, and were instructed to engage in a second interaction that was identical in form to the first. Upon completion of the second interaction, C and P were again separated. C, P, and O completed postmeasures, including the manipulation check, and then were thoroughly debriefed on the purposes of the experiment.

Manipulation and Measure

Cs in the affectionate condition were instructed to "act like you really like your partner" in the second interaction, while Cs in the nonaffectionate condition were told to "act like you really dislike your partner." In both instances, Cs were asked to comply with the instructions using whatever nonverbal behaviors they felt would most naturally communicate liking or disliking to P. That is, Cs were given no instructions on particular behaviors to manipulate, nor were they provided with examples of behaviors that could or should be used. In both conditions, Cs were instructed to begin their manipulations when they were reunited with Ps and to maintain the manipulations throughout the second interaction.

Perceived affection was assessed using a seven-item measure. Four of the items were derived from the affection subscale of the Role Behavior Test (Foa & Foa, 1974) and the remaining three were adapted from the liking scale employed by Floyd and Burgoon (1999).[1]

The Likert-type items had a theoretic range of 1 to 7 with higher scores indicating more affectionate behavior. Cs, Ps, and Os all completed the scale with reference to Cs' behaviors after the first and second interactions for the purpose of checking the manipulation, and Ps' and Os' scores during the postmanipulation portion of the interaction served as the dependent variables in the hypothesis tests. Interitem reliabilities for Cs, Ps, and Os were .75, .73, and .75, respectively.

Analysis of Acoustic Properties

To measure the acoustic features of Cs' vocalic performance during the second, manipulated part of each interaction, we edited the tapes of all 48 sessions into separate segments for each of the participants and each of the confederates. In other words, the edited version of the interactions isolated the speaking turns of each individual and contained only the speech of individual speakers. Audiotapes were then produced and analyzed through the IBM Speechviewer II computer program that computes descriptive statistics (mean and variance) for fundamental frequency (F_0) and acoustic intensity, analyzed ten seconds at a time. These measures were then aggregated for each speaker to produce composite scores for his or her F_0, variance in F_0, and acoustic intensity during the postmanipulation portion of the interaction. During this analysis we also calculated the cumulative duration of speech for each individual in each interaction.

Results

Manipulation Checks

To ensure interactions between strangers, C and P were asked to indicate, on a 7-point scale, how well they knew each other prior to the interaction (with higher scores indicating greater familiarity). Os were given the same 7-item scale to complete in reference to both C and P. Cs indicated that they did not know Ps ($M = 1.77, SD = 1.32$) and Ps indicated that they did not know Cs ($M = 1.74, SD = 1.44$). Likewise, Os indicated that they did not know Ps ($M = 1.12, SD = 0.40$) or Cs ($M = 1.07, SD = 0.35$) prior to the interactions.

To ascertain the success of the affection manipulation, we compared Cs', Ps', and Os' assessments of Cs' affection level in separate mixed-model ANOVAs, with time as the within-subjects factor and behavior manipulation and gender as between-subjects factors. Significant time-by-behavior interactions obtained for Cs' self-reports, $F(1, 44) = 107.45, p < .001$, partial $\eta^2 = .71$; Ps' reports, $F(1, 44) = 21.97, p < .001$, partial $\eta^2 = .33$; and Os' reports, $F(1, 44) = 4.62$, p = .023, partial $\eta^2 = .13$. The means, which appear in Table 1, indicate success for the manipulation according to all three assessments.

TABLE 1
Means and Standard Deviations by Time and Behavior Condition for Manipulation Checks

Measure	Premanipulation		Postmanipulation	
	Hi Aff	Low Aff	Hi Aff	Low Aff
Confederates' self-report	5.67 (0.55)	5.65 (0.70)	6.30 (0.87)	2.95 (1.11)
Participants' report	5.26 (0.60)	5.42 (0.58)	5.55 (1.04)	4.36 (1.19)
Observers' report	4.62 (0.62)	5.44 (0.74)	4.46 (0.69)	4.14 (1.24)

Note. Standard deviations are in parentheses.

Hypotheses

Descriptive statistics for the acoustic measures and talk time are reported in Table 2. (We reported statistics for both the premanipulation and postmanipulation portions of the interaction, even though hypothesis tests were conducted on postmanipulation data only.) For descriptive purposes, we calculated zero-order correlations among the study variables. The coefficients are reported in Table 3. Although the significant correlations between variance in F_0 and Ps' and Os' perceived affection are supportive of the second hypothesis, the correlations involving mean F_0 and mean acoustic intensity, while in the directions predicted, are nonsignificant. However, as Tusing and Dillard (2000) indicated, zero-order correlations do not provide strong tests of the hypotheses because they fail to control for the relationships among independent variables. In fact, some of the independent variables show significant correlations with each other (i.e., r's = .01 to

TABLE 2
Descriptive Statistics for Acoustic Measures

Variable	Unit	Mean	SD	Minimum	Maximum
Men's F_0					
Premanipulation	Hertz	199.96	50.03	50.10	277.80
Postmanipulation	Hertz	204.20	37.04	71.67	252.10
Women's F_0					
Premanipulation	Hertz	203.73	20.97	141.80	240.40
Postmanipulation	Hertz	206.00	19.45	174.50	235.00
F_0 standard deviation					
Premanipulation	Hertz	77.07	37.74	42.70	277.40
Postmanipulation	Hertz	73.07	19.56	43.30	123.00
Acoustic intensity					
Premanipulation	Acoustic energy	5.47	2.51	1.20	11.50
Postmanipulation	Acoustic energy	5.11	2.31	1.00	9.40
Talk time					
Premanipulation	Seconds	88.33	36.28	20.00	180.00
Postmanipulation	Seconds	68.96	37.94	10.00	170.00

TABLE 3
Intercorrelations Among Study Variables

	Variable	I	II	III	IV	V	VI	VII
I.	Receivers' perceived affection	—						
II.	Observers' perceived affection	.45**	—					
III.	Biological sex	.14	.12	—				
IV.	Message length	.06	.20	.34*	—			
V.	F_0 mean	.16	.20	.03	.11	—		
VI.	F_0 variance	.47**	.43*	.36*	.01	.09	—	
VII.	Intensity	−.03	−.04	.36*	.44**	.11	.12	—

Note. Significance tests are two-tailed, based on $N = 48$ interactions. $^{*}p < .05$; $^{**}p < .01$.

.47). As noted below, we dealt with his multicollinearity by standardizing with Z-scores the predictor variables in our regression analyses.

Two hierarchical regressions, with Ps' and Os' perceptions of Cs' affection level as the respective criterion variables, were used to test for hypothesized relationships with F_0, variance in F_0, and acoustic intensity. In both regressions, two control variables were entered in the first two steps. Because numerous studies have indicated that women's communicative patterns are typically more affectionate than men's (for review, see Morman & Floyd, 1998), we entered the sex of the speaker in the first step. In the second step we entered Cs' talk time, or the amount of time each C spoke, to control for its potential influence on judgments of affection. In the third step we entered F_0, variance in F_0 (as a function of its standard deviation), and acoustic intensity. Following Tusing and Dillard (2000), we also included in a fourth step the interaction between biological sex and F_0, due to the biological difference in men's and women's modal F_0 (Titze, 1989).

The results of the regression analyses are reported in Tables 4 and 5. The sex-by-F_0 interaction in the second regression (using Os' perceptions) was nonsignificant. Thus, to produce a more parsimonious model, we reconfigured the second regression by removing the fourth step.

The hypotheses were tested using the beta weights. The first hypothesis predicted a linear relationship between F_0 and perceived affection. For Ps' perceptions, the beta weight for mean F_0 was $-.20$, $t = 1.63$, $p = .13$). However, the sex-by-F_0 interaction term was significant, $\beta = 2.86, t = 2.60, p = .014$. The interaction is plotted in Figure 1, using "high" and "low" groups on F_0 that were configured by using one standard deviation above and one standard deviation below the grand mean. The results of the interaction indicate that, for women, higher mean F_0 is related to greater perceived affection, as

TABLE 4
Hierarchical Regression Predicting Receivers' Perceptions of Affection ($N = 48$)

Predictor Variables	Zero-order *r*	*B*	*SE B*	β	ΔR^2
Step 1 Biological sex	.14	6.67	2.64	2.67*	.86
Step 2 Talk time	.06	.05	.05	1.13	.01
Step 3					
F_0 mean	.16	.05	.02	1.17**	5.35***
F_0 standard deviation	.47**	.03	.01	.55***	
Acoustic intensity mean	−.03	−.04	.08	−.08	
Step 4					
Biological sex-by-F_0	−.15	.03	.01	2.89*	6.79*

Note. Total $R^2 = .39$, adjusted $R^2 = .30$. $F(6,47) = 4.38$, $p = .002$. $^*p < .05$, $^{**}p < .01$, $^{***}p < .005$.

H1a predicts. The opposite pattern obtained for men, however, with lower mean F_0 being associated with greater perceived affection.

For Os' perceptions, the beta weight for mean F_0 was .29, $t = 2.06$, $p = .04$. Thus, H1b was supported.

The second hypothesis predicted a linear relationship between F_0 variance and perceived affection. For Ps' perceptions, the beta weight for F_0 variance was .53, $t = 4.13$, $p < .001$. For Os' perceptions, the beta weight for F_0 variance was .51, $t = 3.38$, $p = .002$. Hypotheses 2a and 2b were therefore supported.

The third hypothesis predicted an inverse relationship between acoustic intensity and perceived affection. For Ps' perceptions, the beta weight for intensity was $-.03$, $t = -.60$, $p = .83$. For Os' perceptions, the beta weight for intensity was $-.21$, $t = -1.31$, $p = .20$. Hypotheses 3a and 3b were not supported.

TABLE 5
Hierarchical Regression Predicting Observers' Perceptions of Affection ($N = 48$)

Predictor Variables	Zero-order *r*	*B*	*SE B*	β	ΔR^2
Step 1 Biological sex	.12	−.13	.20	−.11	.59
Step 2 Talk time	.20	.40	.20	.33*	1.09
Step 3					
F_0 mean	.19	.33	.17	.29*	5.12**
F_0 standard deviation	.43**	.59	.18	.51***	
Acoustic intensity mean	−.04	−.24	.19	−.21	

Note. Total $R^2 = .33$, adjusted $R^2 = .24$. $F(5, 47) = 3.52$, $p = .01$. $^*p < .05$, $^{**}p < .01$, $^{***}p < .005$.

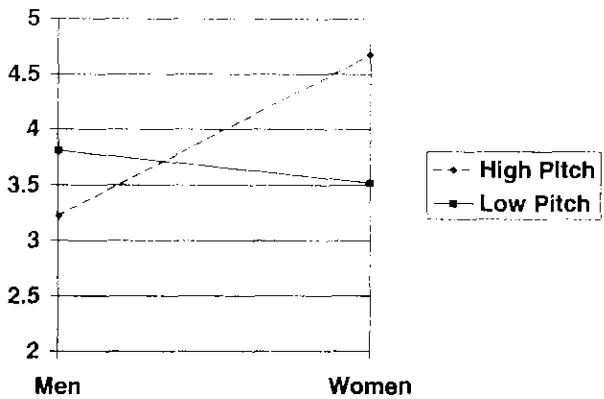

Figure 1. Biological sex-by-F_0 interaction on receivers' perceived affection.

Discussion

In the present investigation we examined relationships between speakers' vocal characteristics and listeners' and observers' perceptions of speakers' affectionate intentions toward listeners. The hypotheses, drawn from affection exchange theory, received moderate support.

Our first hypothesis called for a significant linear relationship between mean F_0 and participants' and observers' perceptions of the confederates' affection levels. A significant beta weight confirmed the prediction for observers' perceptions. However, participants' perceptions were implicated in a sex-by-F_0 interaction term, indicating that while higher mean F_0 was related to greater perceived affection when the confederates (and the participants) were women, lower mean F_0 was related to greater perceived affection for men. Interestingly, this interaction comports with one identified by Tusing and Dillard (2000), in which they found that mean F_0 was linearly related to perceived dominance for men and inversely related to perceived dominance for women. Considered together, these two findings appear to indicate that men's voices are perceived to be more affiliative and less dominant or aggressive when they are lower in pitch, whereas women's voices are perceived to be more affiliative and less dominant when they are higher in pitch.

We are somewhat puzzled by this finding, especially given that men's greater physical size, relative to women, generally gives them the ability to be more physically threatening than women; therefore, AET would predict that it should be men's voices in particular that are characterized as more affectionate when they are higher in pitch (because that should signal smaller physical size and therefore less potential threat). One possible explanation is that, because human men have lower modal F_0 than do women, an excessively high F_0 in a man's voice may signal that the man is weak or effeminate and would therefore not be a good source of relational resources; thus, human evolution has selected against interpreting high male F_0 (but not female F_0) as affectionate or affiliative. A second explanation is that men, to convey liking to other men, relax their adaptive motivation to dominate the interaction. Newton and Burgoon (1990) indicated that lower F_0 was associated with greater relaxation in a conflict situation.[2] We defer investigation of these and other alternative hypotheses to future studies.

The second hypothesis predicted linear relationships between confederates' F_0 variation and participants' and observers' perceptions of the confederates' levels of affection. This prediction was supported for both participants and observers, supporting the argument that because variation in F_0 is perceived to be more pleasant than the lack of variation, humans should associate that pleasantness with a pleasant, affiliative intention on the part of the speaker. Finally, the third hypothesis predicted inverse relationships between confederates' mean vocal intensity and participants' and observers' perceptions of the confederates' levels of affection. Although the beta weights for participants' and observers' judgments were both in the predicted directions, neither achieved significance. This, too, is puzzling in light of the reasoning that louder, more intense vocal sounds should indicate a larger entity than quieter vocal sounds and should therefore communicate greater potential threat.

One possible reason why this pattern did not obtain in the present experiment is that, because of the face-to-face nature of the interactions and the fact that participants and observers had access to the visual channel as well as the verbal and vocal channels, participants and observers were able to ascertain through other means that confederates posed no actual threat to them. As a result, participants and observers may not have taken variations in vocalic intensity into account when forming their impressions of confederates. The fact that the beta weight for intensity corresponding to observers' perceptions was of a greater magnitude than that corresponding to participants' perceptions comports with this alternative explanation, given that participants' physical involvement in the conversations should give them access to more nonverbal information about confederates than was available to observers, who watched the conversations through a closed-circuit television. Future research can address this possibility by holding more sources of nonverbal information constant; of course, this has the disadvantage of reducing ecological validity, given that humans in normal interaction tend to process and respond to multiple nonverbal cues simultaneously, rather than to single cues in isolation (see, e.g., Burgoon, 1994).

A second possible explanation for the failure of Hypothesis 3 is that there is greater variability in relational meaning associated with vocalic intensity than with fundamental frequency, and that such variability overshadowed a main effect of intensity on receivers' and observers' perceptions. For instance, low vocalic intensity might have communicated boredom, disinterest, or apprehension on the part of the speaker instead of an affectionate intention. It may be, in fact, that in the absence of other relevant cues such as gestures or facial expressions, humans do not assign a consistent valence to vocalic intensity. If an association between intensity and perceptions of relational meaning fails to achieve significance in replications, then this would call for reconsideration of this aspect of AET.

Limitations and Conclusions

Some limitations of the current investigation should be borne in mind when interpreting the results. First, the experiment relied on a small, undergraduate sample that, although it is by no means uncommon in research of this type, may limit generalizeability of the results. Second, all of the students in each interaction (confederate, participant, and observer) were of the same sex. This is potentially significant, in that adaptations to interpret vocal behaviors (or other nonverbal behaviors) as affectionate may well affect opposite-sex pairings differently than same-sex pairings, due to reproductive possibilities that are present in the former. For instance, women may assign different interpretations to a man's vocal characteristics than other men do due to adaptive mechanisms that help them to identify potential mates; likewise, men may interpret women's vocal attributes differently than other men's. These hypotheses should be investigated in an experiment that is completely crossed with respect to biological sex. A third limitation is that, although we controlled for some sources of error variance in our regression analyses, other potential sources (such as visual information about confederates; see Burns & Beier, 1973) were uncontrolled. These limitations should temper interpretation of the results but also provide some important avenues for future research.

Despite its limitations, the present study had certain strengths, chief among which is the fact that actual face-to-face interactions were used instead of staged or videotaped stimuli, which others have used (see Floyd, 1999a; Tusing & Dillard, 2000). This is important not only because it more closely approximates the conditions under which individuals routinely make interpersonal judgments about others but also because it allows for greater variance in the vocalic measures than short staged or videotaped segments tend to (the segments used by Tusing & Dillard averaged only 3.5 seconds in length, while those used by Floyd to investigate other nonverbal phenomena averaged approximately 5 seconds in length).

A second strength is that the hypothesis tests were based on acoustic, rather than perceptual, measures of confederates' speech. This frees the measurements from potential bias that is introduced whenever human coders are employed, at least with respect to the generation of the data.[3] A final strength of the present experiment is that it examined the effects of multiple vocal characteristics together, rather than in isolation. This is an important strength because it provides a more realistic approximation of normal, ongoing judgments in face-to-face interaction; that is, people hear and attend to multiple vocal cues at once in normal conversation, so studying vocal cues in isolation can be ecologically invalid (see Street & Brady, 1982).

In summary, the present results provide mixed support for the predictions of AET, with respect to vocalic behavior among previously unacquainted adults. An important extension of this study will be to replicate these tests in other relationship types, to ascertain whether the level of familiarity and/or the nature of the relationship (e.g., whether romantic or platonic) might influence people's tendencies to interpret particular vocal cues as affectionate.

NOTES

1. RBT items were: "I tried to do things that he or she would like," "I ignored my partner's feelings and showed that I didn't like him or her" (reverse-scored), "I showed trust in my partner," and, "I tried to let my partner know that I can't stand him or her" (reverse-scored). Additional items from Floyd and Burgoon (1999) were: "I acted as if I liked my partner," "I made it clear that I was not interested in my partner" (reverse-scored), and, "I seemed to get along well with my partner."
2. We thank one of the reviewers for this observation.
3. Of course, the data are still subject to interpretation within the framework of the theory, and as such, are not completely free of potential bias. By virtue of the process that generated them, however, these hypothesis tests ought to be less subjective than tests employing coders' perceptual measurements. We thank one of the reviewers for this observation, as well.

REFERENCES

Aronovitch, C. D. (1976). The voice of personality: Stereotyped judgments and their relation to voice quality and sex of speaker. *Journal of Social Psychology, 99,* 207–220.

Bombar, M. L., & Littig, L. W. (1996). Babytalk as a communication of intimate attachment: An initial study in adult romances and friendships. *Personal Relationships, 3,* 137–158.

Buller, D. B., & Burgoon, J. K. (1986). The effects of vocalics and nonverbal sensitivity on compliance: A replication and extension. *Human Communication Research, 13,* 126–144.

Burns, K. L., & Beier, E. G. (1973). Significance of vocal and visual channels in the decoding of emotional meaning. *Journal of Communication, 23,* 118–130.

Ferguson, C. A. (1964). Baby talk in six languages. *American Anthropologist, 66,* 103–114.

Ferguson, C. A. (1977). Baby talk as a simplified register. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children* (pp. 209–235). Cambridge, UK: Cambridge University Press.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development, 60,* 1497–1510.

Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development, 64,* 657–674.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20,* 104–113.

Floyd, K. (1999a). All touches are not created equal: Effects of form and duration on observers' perceptions of an embrace. *Journal of Nonverbal Behavior, 23,* 283–299.

Floyd, K. (1999b). To match or not to match: Effects of behavioral congruence on interpersonal connectedness. *Journal of Social Psychology, 139,* 309–322.

Floyd, K. (2001). Human affection exchange: I. Reproductive probability as a predictor of men's affection with their sons. *Journal of Men's Studies, 10,* 39–50.

Floyd, K., & Burgoon, J. K. (1999). Reacting to nonverbal expressions of liking: A test of interaction adaptation theory. *Communication Monographs, 66,* 219–239.

Floyd, K., & Morman, M. T. (1997). Affectionate communication in nonromantic relationships: Influences of communicator, relational, and contextual factors. *Western Journal of Communication, 61,* 279–298.

Floyd, K., & Morman, M. T. (1998). The measurement of affectionate communication. *Communication Quarterly, 46,* 144–162.

Floyd, K., & Morman, M. T. (2000). Affection received from fathers as a predictor of men's affection with their own sons: Tests of the modeling and compensation hypotheses. *Communication Monographs, 67,* 347–361.

Floyd, K., & Morman, M. T. (2001). Human affection exchange: III. Discriminative parental solicitude in men's affectionate communication with their biological and nonbiological sons. *Communication Quarterly, 49,* 310–327.

Floyd, K., & Morman, M. T. (in press). Human affection exchange: II. Effects of role and biological status on affection in father-son relationships. *Journal of Social Psychology.*

Foa, U. G., & Foa, E. B. (1974). *Societal structures of the mind.* Springfield, IL: Charles C. Thomas.

Garnica, O. K. (1977). Some prosodic and paralinguistic features of speech to young children. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children* (pp. 63–88). Cambridge, UK: Cambridge University Press.

Hale, J. L., & Burgoon, J. K. (1984). Models of reactions to changes in nonverbal immediacy. *Journal of Nonverbal Behavior, 8,* 287–315.

Harrigan, J. A., Gramata, J. F., Lucic, K. S., & Margolis, C. (1989). It's how you say it: Physicians' vocal behavior. *Social Science and Medicine, 28,* 87–92.

Kimble, C. E., Forte, R. A., & Yoshikawa, J. C. (1981). Nonverbal concomitants of enacted emotional intensity and positivity: Visual and vocalic behavior. *Journal of Personality, 49,* 271–283.

Morman, M. T., & Floyd, K. (1998). "I love you, man:" Overt expressions of affection in male-male interaction. *Sex Roles, 38,* 871–881.

Ohala, J. J. (1982). The voice of dominance. *Journal of the Acoustical Society of America, 72,* S66.

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F_0 of voice. *Phonetica, 41,* 1–16.

Owen, W. F. (1987). The verbal expression of love by women and men as a critical communication event in personal relationships. *Women's Studies in Communication, 10,* 15–24.

Ray, G. B. (1986). Vocally cued personality prototypes: An implicit personality theory approach. *Communication Monographs, 53,* 266–276.

Salt, R. F. (1991). Affectionate touch between fathers and preadolescent sons. *Journal of Marriage and the Family, 53,* 545–554.

Shute, B., & Wheldall, K. (1989). Pitch alterations in British motherese: Some preliminary acoustic data. *Journal of Child Language, 16,* 503–512.

Street, R. L., & Brady, R. M. (1982). Speech rate acceptance ranges as a function of evaluative domain, listener speech rate, and communication context. *Communication Monographs, 49,* 290–308.

Titze, I. R. (1989). Physiologic and acoustic differences between male and female voices. *Journal of the Acoustical Society of America, 85,* 1699–1701.

Toda, S., Fogel, A., & Kawai, M. (1990). Maternal speech to three-month-old infants in the United States and Japan. *Journal of Child Language, 17,* 279–294.

Tusing, K. J., & Dillard, J. P. (2000). The sounds of dominance: Vocal precursors of perceived dominance during interpersonal influence. *Human Communication Research. 26.* 148–171.

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responses. *Canadian Journal of Psychology, 43,* 230–246.

White, C. H., & Burgoon, J. K. (1997, November). *Adaptation and communicative design: Patterns of interaction in truthful and deceptive conversations.* Paper presented to National Communication Association, Chicago, IL.

Zebrowitz, L. A., Brownlow, S., & Olson, K. (1992). Baby talk to the babyfaced. *Journal of Nonverbal Behavior, 16,* 143–158.