



The siren song of vocal fundamental frequency for romantic relationships

Sarah Weusthoff¹*, Brian R. Baucom² and Kurt Hahlweg¹

¹ Clinical Psychology, Psychotherapy, and Assessment, Department of Psychology, Technische Universität Braunschweig, Braunschweig, Germany

² Department of Psychology, University of Utah, Salt Lake City, UT, USA

Edited by:

Anjali Bhatara, Université Paris Descartes, France

Reviewed by:

Alexandra Suppes, Columbia University, USA
Cheryl Carmichael, City University of New York, USA

*Correspondence:

Sarah Weusthoff, Clinical Psychology, Psychotherapy, and Assessment, Department of Psychology, Technische Universität Braunschweig, Humboldtstrasse 33, 38106 Braunschweig, Germany
e-mail: s.weusthoff@tu-bs.de

A multitude of factors contribute to why and how romantic relationships are formed as well as whether they ultimately succeed or fail. Drawing on evolutionary models of attraction and speech production as well as integrative models of relationship functioning, this review argues that paralinguistic cues (more specifically the fundamental frequency of the voice) that are initially a strong source of attraction also increase couples' risk for relationship failure. Conceptual similarities and differences between the multiple operationalizations and interpretations of vocal fundamental frequency are discussed and guidelines are presented for understanding both convergent and non-convergent findings. Implications for clinical practice and future research are discussed.

Keywords: fundamental frequency, attraction, emotional arousal, romantic relationships, couples, communication, conflict

INTRODUCTION

Across cultures, adults rate marriage and long-term monogamous relationships as the most important ones in their lives (Buss, 2005). Despite this, divorce rates in industrialized countries rank between 40 to 55% (Hahlweg et al., 2010). A tremendous amount of effort has been accordingly devoted to identifying and understanding risk factors for relationship dissolution (Gottman and Notarius, 2002) with a particular emphasis on identifying risk factors that are present from the outset of a relationship (Caughlin and Huston, 2002). The polarization model of romantic relationships (Jacobson and Christensen, 1998; Baucom and Atkins, 2013) suggests that one of the most important kinds of risk factors originates in variables that are initially associated with attraction and desire that later become sources of distress and interpersonal friction. A growing body of empirical evidence suggests that the fundamental frequency of romantic partner's voices may represent precisely this kind of risk. Vocal qualities that are found to be attractive during the early stages of relationship formation are also associated with increased levels of dysfunction, increased risk for divorce, and decreased likelihood of benefitting from couple therapy in later stages of a relationship.

Evolutionary models suggest that perceptions of health status and likelihood of reproductive success are one of the primary determinants of attractiveness and mate selection. Physical aspects of attractiveness like facial symmetry, body height, and the voice are indicative of levels of sexual hormones, and thus of the likelihood of producing healthy offspring. Generally, indicators of higher levels of masculinity in males (based on higher levels of testosterone) and higher levels of femininity in females (based on higher levels of estrogen) are perceived as more attractive and are

linked to better reproductive success. This leads to a dimorphism in attractiveness ratings: males tend to prefer female partners with higher f_0 voices and smaller body sizes while female tend to prefer males with lower f_0 voices and larger body sizes (Puts et al., 2012). Though these sex-related differences are important determinants of attraction, the polarization model of relationship distress suggests that they may become associated with dysfunction and increased risk for divorce as relationships mature. A change from attraction to distress is often accompanied by a change in attribution for the difference itself. Differences that are seen as enriching and complimentary are often experienced as desirable (Aron and Aron, 1997) while differences that are seen as shortcomings and faults in the other commonly contribute to a cycle of distress. When differences are seen as faults in the other, spouses typically alternate between criticizing and blaming one another for the problems in their relationship and defending themselves from their partner's attacks (i.e., they engage in a strong demand/withdraw cycle of conflict). As this process unfolds, the relationship becomes polarized, and the three core aspects of interaction, communication, perception, and physiology, become imbalanced (Burman and Margolin, 1992; Gottman, 1993), conflict ensues, and even routine interaction with the spouse becomes highly aversive. The aversive nature of this cycle typically results in increasing polarization over time and makes it very difficult for spouses to resolve even minor problems. Being stuck in the polarization cycle is highly distressing for spouses, and high levels of aversive arousal during interaction are one of the most robust predictors of risk for relationship distress and dissolution (Gottman and Levenson, 1992). Aversive arousal has most often been captured via well-established physiological indices like heart rate (HR), blood pressure (BP), or skin conductance

(Larsen et al., 2008) as well as via endocrine measures like cortisol, or epinephrine (Robles and Kiecolt-Glaser, 2003; Ditzen et al., 2011). In addition to well-replicated findings based on physiological measures, a growing body of evidence shows that higher levels of vocally encoded emotional arousal, which are reflected in higher levels of f_0 , are similarly related to increased risk for a wide range of negative relationship outcomes.

FUNDAMENTAL FREQUENCY (f_0)

As a mathematical quantity, fundamental frequency (f_0) refers to the lowest frequency harmonic of the speech sound wave. This frequency is biologically determined by the pattern of vibration created by the vocal folds during phonation, the phase of speech production where the outward flow of air from the lungs is regulated by the larynx. Higher rates of opening and closing of the vocal folds across the glottis are associated with higher f_0 values (measured in cycles per second or Hertz [Hz]). Perceptually, f_0 is highly correlated with pitch where higher f_0 values correspond to higher pitch (Juslin and Scherer, 2005). F_0 can be easily calculated using Praat, a freely available, Windows-based software package (Boersma and Weenink, 2013; www.praat.org), to analyze existing audio recordings of speech in adequate quality. F_0 is assessed continuously during human speech and can change rapidly. Different parameters like mean f_0 , f_0 range, minimum f_0 , or maximum f_0 can be calculated and used in empirical research. All of those can be calculated either at a very small scale (such as for each talk turn) or averaged across over the entirety of a conversation. As is true of much behavioral research, there is greater agreement about how to calculate f_0 than there is about how to understand what f_0 represents. F_0 has been variously interpreted as an index of vocally encoded emotional arousal (see Juslin and Scherer, 2005 for a review; Weusthoff et al., 2013) and dominance (Puts et al., 2006; Borkowska and Pawlowski, 2011), and additional work demonstrates that f_0 correlates with other factors such as age and pubertal (Hollien and Shipp, 1972; Brown et al., 1991; Hollien et al., 1994, 1997), phonemic and syntactic structure of speech (Whalen and Levitt, 1995). Recent work examining f_0 during social interaction provides a framework for integrating the various interpretations of f_0 . In Weusthoff et al.'s 2013 examination of simultaneous associations between f_0 , biological sex, physiological indices of arousal, and social behavior, biological sex was the largest predictor of individual differences in f_0 while physiological arousal and social behaviors were the best predictors of variance in f_0 attributable to a specific social interaction. These results suggest that f_0 can be understood as conveying information about both traits (such as biological sex) and states (momentary physiological arousal and social behavior) and highlights the need for careful analysis of f_0 to allow for specific interpretations.

METHOD OF REVIEW

In contrast to the nascent body of research examining f_0 during marital interactions, f_0 has been intensively studied using different research paradigms across multiple disciplines (e.g., single conversations recorded during naturally occurring stressful events like the New York City blackout in 1977, Streeter et al., 1982, or emotion portrayals by professional actors, Banse and Scherer, 1996; see Scherer, 2003 for a review). For example, a key-word based

search on SCOPUS (using the search terms “social interaction” AND “verbal” OR “non-verbal” in “Keywords, Abstract, or Title”) yielded more than 2000 results. Here, we focus on research with a specific focus on relationship formation and maintenance. The body of work on f_0 and attraction is robust and well-developed. Thus, we seek to provide a representative sampling of the main findings in this area. In contrast, there are many fewer studies of naturalistic social interaction in general much less about marital interaction specifically. As others have noted, studies using naturalistic vocal expressions of affect with sufficient intensity of the emotions displayed are needed in vocal expression research in order to obtain speech that offers sufficient levels of “ecological validity” (Juslin and Scherer, 2005). Communication in close relationships seems to fulfill these criteria but research has only recently begun to take f_0 into account here.

Research with human participants can be split up in two further groups: whether interlocutors had some form of relationship with each other before interacting, or not. In the case of pre-existing relationship, four main areas of relationships have been covered: the ones between parents and infants (see Irwin, 2003 for a review), between psychotherapists and clients (see Greenberg and Pascual-Leone, 2006 for a review), between physicians and patients (see Hassan et al., 2007; Hulsman et al., 2011 for a review), and between spouses in intimate relationship. Independent of targets, f_0 has been found to be associated with arousal. As it is beyond the scope of this review to cover this all, the interested reader is referred to the according reviews. In couples interactions, however, the role of vocally encoded emotional arousal has not been investigated closely. This review aims at looking into f_0 and its associations with couple functioning more closely. The authors identified a total number of $N = 5$ studies for this task. All analyses were based on dyadic communication settings (conflict discussions) between spouses from long-term or married heterosexual couples. Instructions for the tasks were standardized, and recordings of the interactions were videotaped during assessment sessions conducted in research laboratories. Participants were either English (Baucom et al., 2009, 2011), or German (Baucom et al., 2012b; Kliem et al., 2012; Weusthoff et al., 2013) native speakers (for a detailed description of the studies included in this review, please see Table 1).

As also described in Weusthoff et al. (2013), f_0 was obtained continuously from the speech of each person during the problem discussion of the assessment point(s). Conversations were segmented per speaker using audio editing programs like Adobe Premier Pro, resulting in data per speaker containing only conversation parts during which he or she talked separately without any other human, or background sound being present. Bandpass filtering was applied to all voice samples prior to further analyses (calculating f_0 values) in order to restrict f_0 values to the normal range of emotional adult speech. The typical range for male speakers is 75–150Hz, for female speakers 150–300Hz, though wider limits can be observed during highly aroused emotional states (Owren and Bachorowski, 2007). Common f_0 mean values are around 225 Hz for women, and about 120 Hz for men. Across the lifespan, women's mean f_0 decreases while male speakers' mean f_0 scores also decrease first, but start to rise again. Changes in adults are most prominent for both sexes between age 50 and

Table 1 | Methodological details of interaction studies reviewed.

Study	participants and communication setting	design and independent measures (covariates)	dependent measures	relevant findings
Baucom et al. (2009)	130 heterosexual, US-American couples in a couple therapy study; conflict discussion with spouse in laboratory (10 min)	longitudinal; sociodemographics, intrapersonal predictors, communication behavior, interpersonal predictors; all assessed at pre-treatment	response to treatment 2 years after treatment termination (deteriorated, unchanged, improved, recovered)	Across therapies, lower levels of wives' f ₀ range linked to a higher likelihood for being in recovered category at the two-year follow-up.
Baucom et al. (2011)	heterosexual, US-American couples (n = 130 in a couple therapy study, n = 38 distressed community couples); conflict discussion with spouse in laboratory (10 min)	cross-sectional; mean f ₀ , relationship satisfaction, influence tactics	demand/withdraw behavior and role	Higher levels of demand/withdraw-behavior were independently linked to higher levels of mean f ₀ across gender. Demanders experienced higher levels of mean f ₀ , compared to withdrawers.
Baucom et al. (2012b)	49 heterosexual, German couples participating in couple-relationship education (CRE); conflict discussion with spouse in laboratory (15 min)	longitudinal; f ₀ range (biological sex, relationship satisfaction)	number of skills taught in the program remembered 11 years after participation in CRE by spouses	Higher levels of pre-training f ₀ range were associated with fewer skills remembered 11 years after the training. Women remembered more skills than men.
Weusthoff et al. (2013)	67 heterosexual, German couples participating in CRE; conflict discussion with spouse in laboratory (15 min)	cross-sectional; psychophysiology, communication behavior, biological sex	f ₀ range	Higher levels of f ₀ range significantly linked to higher levels of physiological arousal, more negative and less positive communication behavior, and being female
Kliem et al. (2012)	68 heterosexual, German couples participating in CRE; conflict discussion with spouse in laboratory (15 min)	longitudinal; f ₀ range, relationship satisfaction, psychophysiology, communication behavior, depressive symptoms	relationship status 11 years after participation in CRE	For females, higher levels of f ₀ range significantly linked to a higher risk for relationship dissolution. For males, higher levels of diastolic blood pressure linked to higher risk for relationship dissolution.

60, with women additionally experiencing hormonal influences associated with the onset of the menopause (Hollien and Shipp, 1972; Brown et al., 1991; Hollien et al., 1997). Though it is possible for human speakers to generate sound outside the typical filtering limits of 75 and 300 Hz, f₀ scores outside of this range are very likely resulting from background machine or electronic noise. Minimum, maximum, and mean f₀ values were generated separately for each spouse by analyzing their respective segmented audio recordings using Praat, a free multiple platform program (Boersma and Weenink, 2013). F₀ range was calculated separately for each partner at each assessment by subtracting the partner's averaged minimum f₀ from the partner's averaged maximum f₀ across the whole conversation. In Baucom et al.'s (2011) study, f₀ mean was derived from raw f₀ scores across the whole conversation. For a more detailed description of potential f₀ range indices, please see **Table 1** in Baucom et al. (2012a).

f₀ IN INTIMATE RELATIONSHIP RESEARCH

Due to three reasons, it is important to further investigate the role of f₀ in intimate relationships. One, relationship deterioration and/or separation are associated with tremendous costs for affected individuals and society as a whole, both concurrently and in the long run. Dysfunctional relationships lead to poorer mental and physical health in spouses (Kiecolt-Glaser and Newton, 2001), a higher likelihood for divorce and reduced levels of social support (Amato, 2010). Due to higher levels of unhealthy behaviors like smoking, even a decreased life expectancy (Larson and Halfon, 2013) in affected children has been observed. Second, couple-relationship education programs (CRE) aim at reducing the risk of relationship dissolution by teaching communication skills and ways of dealing with aspects of couple's emotional lives, e.g., *EPL – Ein Partnerschaftliches Lernprogramm* ("A Learning Program for Couples"; Kaiser et al., 1998), especially during conflict discussions. These programs are known to be highly effective in doing so both in the short and in the long run (Hawkins et al., 2008; Hahlweg and Richter, 2010). However, the mechanisms of change in CRE and how emotional arousal affects the work in and outcome of CRE are still unknown. Communication is assumed to be the central mechanism of change, however, empirical studies using the paradigm of couples having videotaped conflict discussions that are later rated and / or analyzed with regard to different aspects of behavior have not consistently supported this supposition. One reason for this might be the rather simplistic ways of operationalizing a complex behavior like human communication as single elements like positive, or negative utterances (Christensen, 2010). Therefore, research should consider several aspects of communication simultaneously in a multi-channel approach (Gottman and Notarius, 2002). Third, though emotional arousal is known to play a crucial role in relationship dissolution (Gottman and Levenson, 1992), the specific pathways and interactions with other important variables like communication behavior, or relationship satisfaction are still unknown (Weusthoff et al., 2013). As future research in this field should focus on how to "modify emotion-driven, dysfunctional, and destructive interactional behavior," and to "elicit avoided, emotion-based" behavior (Christensen, 2010, pp. 36-37), there is additional need for an objective tool for detecting and

explaining emotional arousal in the context of couple interactions and intimate relationships.

During courtship, f₀ is one of the evolutionary most important signals for non-visual gender discrimination (which is important for successful identification of potential mates; Junger et al., 2013), and for judging the attractiveness of a potential spouse (Borkowska and Pawlowski, 2011). Very masculine voices in males (perceived as low voice pitch, and thus low f₀), and very feminine voices in females (perceived as high voice pitch, and thus high f₀) are perceived as particularly attractive (DeBruine et al., 2006; Feinberg et al., 2006, 2012; Borkowska and Pawlowski, 2011; O'Connor et al., 2012). These perceptions are attenuated in both men and women during high-fertile phases of the female menstrual cycle (Puts, 2005; Pipitone and Gallup, 2008; Hodges-Simeon et al., 2010). High vocal masculinity in males are linked to high levels of testosterone, and associated with a higher likelihood for producing healthy offspring (O'Connor et al., 2012). However, in the long run, males with higher levels of vocal masculinity report a larger number of different sexual partners, and are less likely to engage in relationship maintaining and parental behaviors (O'Connor et al., 2011, 2012). It also seems plausible that f₀ during relationship initiation stages displays high levels of arousal that is taken by partner as an indicator for higher engagement in a relationship, also hinting at a process similar to polarization.

During maintenance, less research on f₀'s role has been conducted so far. F₀ has most often been studied in couples' conflict interactions where it seems to index levels of emotional arousal. F₀ during couple conflict has been demonstrated to be significantly positively associated with multiple cardiac, and endocrine indices of arousal, and linked to more negative and less positive communication behavior. F₀ is thought to simultaneously display autonomic physiological as well as socially learned reactions in one signal (Weusthoff et al., 2013). Furthermore, it has been shown that spouses influence each other in their levels of arousal. If one partner is highly aroused, it becomes more difficult for the other one to maintain on a functional level of arousal, thus leading to polarization processes with regard to f₀ (Baucom and Atkins, 2013).

As also noted elsewhere (Weusthoff et al., 2013), f₀ offers a number of methodological and conceptual benefits with regard to being further used in research on romantic relationships. Its non-invasive nature and lack of need for additional equipment (only a good audio recording device is needed) makes the assessment of f₀ an excellent candidate for data collection in situations like conflict interactions that are most often videotaped, or for post-hoc analyses (e.g., spousal discussions, or therapy sessions). Additionally, f₀ can be computed and analyzed from the videotapes even after conducting a study in which it was not of primary interest. Furthermore, f₀ is more closely associated with individual psychological than physical distress (Johannes et al., 2007), making it more sensitive to detect changes in psychological load and less prone to artifacts that influence physiological measures of arousal like HR or skin conductance (e.g., movements, Sloan and Kring, 2007). Being based on physical properties of speech, f₀ can be analyzed objectively and independent of the researcher's native language and culture (Weusthoff et al., 2013). Perhaps most importantly, f₀ is directly involved in and available during the

process of communication and can be perceived by a listener (Baucom et al., 2011). There is good evidence that partners in communication directly respond to each other's f₀ without being aware of this fact (Gregory and Webster, 1996). Similar to cardiovascular indices of arousal (e.g., HR), f₀ seems to encode aspects of emotional arousal that are part of an autonomic process not entirely controllable by a speaker, similar to conditioned emotional responses (Kliem et al., 2012).

Merging these theories, expression of emotions (emotional expression) toward one's spouse can be considered as a central part of well-functioning intimate relationships, the role of underlying emotional arousal, however, seems to be somewhat different. Endocrine and psychophysiological indices of heightened levels of emotional arousal have consistently been linked to unwanted relationship outcomes like dysfunctional communication behavior, higher risk for divorce and separation, or lower levels of relationship satisfaction (Gottman and Levenson, 1992; Kiecolt-Glaser and Newton, 2001; Gottman and Notarius, 2002). As emotions and their underlying processes like arousal are essential parts of social interactions (Juslin and Laukka, 2003) and can influence them heavily (Scherer, 2005), it makes them important information not only for researchers in the social interaction area but also for people's everyday lives.

STUDIES REVIEWED

SUMMARY OF FINDINGS

Across studies, languages, and gender, f₀ has been treated as an index of emotional arousal. Like other indices f₀ has been associated with a number of different negative aspects of couple functioning both concurrently and in the long run. Significant associations have been found in conflict interactions of distressed couples, either participating in couples therapy, or in CRE. More specifically, higher levels of a person's physiological indicators of emotional arousal like HR, BP, or salivary cortisol were associated with higher levels of one's own f₀. Higher levels of a spouse's f₀ have also been linked to higher levels of negative communication behavior, both observed and self-reported, and to lower levels of observed non-verbal positive communication behavior. In highly distressed couples, the likelihood for long-term success of couple therapy (being in the recovered range 2 years after treatment termination) was higher when wives' displayed lower levels of f₀ in a conflict discussion prior to treatment. Elevated levels of f₀ in conflict discussions prior to participation in CRE also lead to a higher likelihood for separation and divorce, and a smaller number of communication skills remembered 11 years after the CRE.

DISCUSSION

This review aimed at examining research on vocally encoded emotional arousal, namely f₀, in close personal relationship communication (interaction between spouses). Significant associations were found with psychophysiological and endocrine indices of emotional arousal, observed and self-reported communication behavior, and relationship functioning (stability and satisfaction). Furthermore, f₀ has been found to be related to skills remembrance in CRE. Significant results were found for concurrent as well as longitudinal links between f₀ and variables of interest, and for a time frame as long as 11 years. Significant gender differences

emerged for f₀ range's association with some variables of interest but not for f₀ mean's.

FINDINGS CONSIDERED FROM AN EVOLUTIONARY PERSPECTIVE

Across studies, f₀ was analyzed as an index of emotional arousal thought to display information about the internal state of the speaker to an interaction partner via the human voice. Stable associations between f₀ and different forms of communication behavior (positive, negative, observed, self-reported, verbal, and non-verbal) emerged across studies. Social signaling theories assume evolutionary reasons for this: In order to heighten chances of survival in ancient hunter-gatherer societies, emotions were adaptive developments helping to avoid danger and to cooperate with others (Buss, 2005). Vocally encoded emotional arousal enables individuals to communicate emotions non-verbally from one person to another, independent of words and language. Vocal communication of emotion is more likely to appear in goal-relevant behavior (Juslin and Laukka, 2003). F₀ scores indexing vocally encoded emotional arousal should therefore be higher during goal-relevant behavior. What is considered as goal-relevant behavior depends on situational aspects: Couple and family conflict is conceptually thought of as a chronic stressor (Doss et al., 2004) where often a wide range of negative communication behavior (like DW) is displayed in order to achieve change in one's spouse (Christensen et al., 2006) and high levels of emotional arousal are displayed via multiple channels (Kiecolt-Glaser and Newton, 2001). Positive associations between f₀ and these variables as found in the reviewed studies investigating conflict could be indexing goal-achieving strategies by spouses.

With regard to the theoretical foundation of emotional arousal, coordinated responses in multiple physiological systems stem from the same source, namely the cognitive appraisal in a given situation and the following changes in autonomous and somatic nervous system activity (Scherer, 2009). These changes were assessed via multiple channel of arousal in Weusthoff et al.'s 2013 study, with the associations between f₀ and psychophysiological indices of arousal being in expected directions. With the periaqueductal grey (PAG), a brain-stem based areal (and thus an evolutionary old one) in humans is involved in both vocal and cardiovascular responses to different kinds of stress. The PAG is thought to integrate emotional aspect of stress responses into the autonomic nervous system responses in different channels (Linnman et al., 2012). These findings suggest that f₀ as an index of emotional arousal, and thus the intensity of the emotional reaction (Weusthoff et al., 2013), are influenced by both basic biological processes as well as by socially learned communication behaviors stemming from a similar evolutionary basis (Juslin and Laukka, 2003).

Different brain regions are found to be involved in processing male and female voices (Sokhi et al., 2005), and discriminatory performance for human voices was better for opposite-sex stimuli than for same sex-stimuli (faster identification; Junger et al., 2013). F₀ seems to enable a stronger attendance toward stimuli higher in social significance like potential mates (Feinberg, 2008), which could be considered as an important aspect of relationship functioning not only in initiation but also in maintaining phases.

The studies covered in this review seem to foster this interpretation: during later phases of relationship and especially in “rough” times, spouses seem to especially use f₀ as a non-verbal signal in relationship-relevant information like expression of distress.

FINDINGS WITH REGARD TO GENDER DIFFERENCES

Across age and interaction settings, gender differences were found in a number of studies. Biologically speaking, f₀ is based on vocal cord length and tension. Given the physiological differences in body size and thus, throat length between men and women, gender differences in f₀ emerge after puberty with females on average having significantly higher f₀ scores than males (Titze, 1989). Two studies in this review (Baucom et al., 2012b; Weusthoff et al., 2013) explicitly investigated gender differences and as expected found higher f₀ range scores for female than for male speakers. Except for Baucom et al. (2011) and Weusthoff et al. (2013), all studies investigating adult speakers included in this review found variables of interest to have significant associations with f₀ range for female speakers only, or found different associations for men and women. Given that female partners are more likely to seek help with regard to problems and/or conflict in close personal relationships (O’Brien, 1988) it seems also plausible that gender differences could stem from different goals and resulting goal-oriented behavior in males and females (especially as no associations being significant for men only emerged). Finding the best partner to produce healthy offspring seems to be an important goal during selection of a potential mate and sex partner during courtship. Significant associations between perceived, and self-rated attractiveness and health, and f₀ for both sexes during sexual selection (though due to endocrine reasons in different directions, with higher attractiveness and perceived health being associated to f₀ scores in women and lower f₀ scores in men; Puts, 2010) also hint at f₀ being an important aspect in goal-relevant behavior.

LIMITATIONS

The work reviewed in this manuscript has taken into account only three indicators of vocally encoded emotional arousal out of a wide variety of potential ones (Juslin and Scherer, 2005, pp. 103–104): f₀ range, f₀ mean, and f₀-time-to-peak as a time-varying aspect of f₀. This has happened for a number of reasons. As earlier work on vocally encoded emotional arousal has focused on f₀ mean in a number of different research settings (i.e., f₀ in speech of psychiatric patients, Tolkmitt et al., 1982; or emotion portrayal by professional actors, Banse and Scherer, 1996), it was chosen as the parameter of interest in the first study conducted on f₀ in close personal relationships (Baucom et al., 2011) in order to enable comparisons to these published empirical findings. However, among the different f₀ indices that can be calculated f₀ range seems to bear a number of advantages in research on close personal relationships (see Juslin and Scherer, 2005 for a review), and was therefore chosen for the studies conducted later. f₀ range is considered to be the cleanest vocal indicator of emotional arousal, and to depict the biggest amount of information on emotional arousal (Juslin and Scherer, 2005; Busso et al., 2009). Furthermore, it facilitates the interpretation of gender differences as it adjusts for potential individual differences by way of its calculation (subtracting the individual’s minimum f₀ score from the individual’s

f₀ maximum score; Weusthoff et al., 2013). Nonetheless, it is possible that other indicators of vocally emotional arousal (e.g., jitter) might contain additional information on the nature of emotional arousal in close personal relationship communication.

Though biological reasons for differences in male and female voices are well-documented and have been discussed in this review, it remains unclear why expected sex differences for f₀ indices and their associations with different variables of interest did not emerge consistently across studies. Particularly, none of the studies found significant gender differences in the associations between f₀ mean (compared to studies using f₀ range) and variables of interest. Future research should regard gender as a covariate in f₀ research in close personal relationship but more detailed investigations on the details of gender differences in f₀ indices are needed.

Investigating emotional arousal in human speech using f₀ includes the drawback of participants being able to willingly influence and control f₀ to a certain degree. Empirical evidence has shown that conscious manipulations can result in both elevated and lowered pitch levels that are identifiable by a communication partner (Kuenzel, 2000). However, given the arousing, stressful, and cognitively demanding situation of couple conflict, it is quite unlikely that participants in the studies reviewed made use of this hypothetical possibility.

IMPLICATIONS FOR CLINICAL WORK AND FUTURE RESEARCH

As f₀ seems to be able to display an individual’s internal levels of emotional arousal, other dyadic settings in which emotional arousal is considered to be important information could also benefit from research on vocally encoded arousal. During social support interactions between spouses, showing the interaction partner positive and helpful behaviors in order to help him or her to deal with stressful situations is the main goal. Social support is considered to consist of different positive and helpful behaviors being observable in communication between interaction partners leading to changes in individual cardiovascular, and endocrine functioning which influences individual health outcomes. Consistent with this model, lower levels of social support behavior have been empirically linked to higher levels of physiological indices of arousal (most often BP), and to poorer health outcomes in affected individuals (Uchino et al., 1996). Given the associations between BP, and HR with f₀ in couple conflict, vocally encoded emotional arousal also seems to be important during social support interactions between spouses. However, in comparison to conflict discussions, more positive and less emotionally arousing behaviors seem to be goal-relevant in social support and should therefore be displayed more often between interaction partners (Verhofstadt et al., 2005). Links between f₀ and variables of interest in social support behavior could therefore be of different direction and magnitude than the ones found in conflict discussions.

Communication Accommodation Theory (CAT; Giles and Coupland, 1991) describes changes in partners’ communication styles during a conversation leading to more, or less similarity between the partners (convergence and divergence). Convergence can occur on different levels (e.g., pitch pattern, speech rate, or emotional expression), and has empirically been shown to occur in

various dyads. Convergence seems to happen somewhat naturally but is associated with different outcomes depending on the interaction setting and other external factors in intimate relationships. Convergence seems to be a beneficial process during interactions asking for social support between partners. For example, high levels of emotional convergence (meaning similarity in emotional expressions between spouses) while sharing emotional events of one's day are related to higher levels of concurrent and longitudinal relationship satisfaction, and higher relationship stability (Anderson et al., 2003). During couple conflict, however, convergence in emotional expression and behavior like the DW pattern have negative impacts on relationship functioning and outcomes (Lee et al., 2012; Baucom and Atkins, 2013). A closer look at potential reasons for these differential effects of convergence in couples' communication could be a fruitful avenue for future research. Are the associations due to levels of relationship satisfaction (happy vs. unhappy couples), do they depend on the context of the interaction (social support vs. conflict), or do both factors and/or a third one contribute?

Vocal characteristics between talkshow host and guests have also been demonstrated to converge across the course of an interview. However, across different dyads, status and/or power seem to be an important factor influencing convergence: The conversation partner lower in status and power (having a greater "need for social approval"; Giles and Coupland, 1991, p. 73) exhibits more change in vocal and emotional expressions, moving in

the direction of the more powerful partner (Gregory and Webster, 1996).

Power processes are also influential in couples communication (Baucom et al., 2011), and are interdependently linked to each other. For example, f₀ in distressed spouses' problem-solving discussions is known to converge in terms of magnitude: if one spouse is highly aroused, the other partner is also more likely to be highly aroused. Furthermore, this covariation seems to lead to problems in partner's ability to regulate their own state of arousal to a comfortable level (Baucom and Atkins, 2013).

Given these findings, it seems likely that spouses' f₀ scores could also be associated with each other across speakers. Gottman and Notarius (2002, p.185) explicitly state that there is a "need for continued focus on sequences or patterns of interaction" in order to identify beneficial and harmful aspects of spousal communication. Sequential analyses of f₀ could shed light on presence, magnitude, and direction of interdependent and cyclical aspects of emotional arousal in couple communication.

ACKNOWLEDGMENTS

This research was supported by grants from the Deutsche Forschungsgemeinschaft (DFG; DFG Fe 263/5-1, Ha 1400/4-1, Ha 1400/16-1, and Ha 1400/16-2) awarded to Kurt Hahlweg and from the National Institute of Child Health and Human Development (F32 HD060410) and the University of Utah awarded to Brian R. Baucom.

REFERENCES

- Amato, P. R. (2010). Research on divorce: continuing trends and new developments. *J. Marriage Fam.* 72, 650–666. doi: 10.1111/j.1741-3737.2010.00723.x
- Anderson, C., Keltner, D., and John, O. P. (2003). Emotional convergence between people over time. *J. Pers. Soc. Psychol.* 84, 1054–1068. doi: 10.1037/0022-3514.84.5.1054
- Aron, A., and Aron, E. N. (1997). "Self-expansion motivation and including other in the self," in *Handbook of Personal Relationships: Theory, Research, and Interventions*, ed. S. Duck (London: Wiley), 251–270.
- Banse, R., and Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* 70, 614–636. doi: 10.1037/0022-3514.70.3.614
- Baucom, B. R., and Atkins, D. C. (2013). "Polarization in marriage," in *Family Theories: A content-based approach*, eds M. Fine and F. Fincham (New York: Routledge), 145–166.
- Baucom, B. R., Atkins, D. C., Eldridge, K., McFarland, P., Sevier, M., and Christensen, A. (2011). The language of demand/withdraw: verbal and vocal expression in dyadic interactions. *J. Fam. Psychol.* 25, 570–580. doi: 10.1037/a0024064
- Baucom, B. R., Atkins, D. C., Simpson, L., and Christensen, A. (2009). Prediction of response to treatment in a randomized clinical trial of couple therapy: a 2-year follow-up. *J. Consult. Clin. Psychol.* 77, 160–173. doi: 10.1037/a0014405
- Baucom, B. R., Saxbe, D. E., Ramos, M. C., Spies, L. A., Iturralde, E., Duman, S., et al. (2012a). Correlates and characteristics of adolescents' encoded emotional arousal during family conflict. *Emotion* 12, 1281–1291. doi: 10.1037/a0028872
- Baucom, B. R., Weusthoff, S., Atkins, D. C., and Hahlweg, K. (2012b). Greater emotional arousal predicts poorer long-term memory of communication skills in couples. *Behav. Res. Ther.* 50, 442–447. doi: 10.1016/j.brat.2012.03.010
- Boersma, P., and Weenink, D. (2013). *Praat: Doing Phonetics by Computer [Computer Program]*, Version 5.3.42. Available at: from <http://www.praat.org> (retrieved March 22, 2013).
- Borkowska, B., and Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Anim. Behav.* 82, 55–59. doi: 10.1016/j.anbehav.2011.03.024
- Brown, W. S. Jr., Morris, R. J., Hollien, H., and Howell, E. (1991). Speaking fundamental frequency characteristics as a function of age and professional singing. *J. Voice* 5, 310–315. doi: 10.1016/S0892-1997(05)80061-X
- Burman, B., and Margolin, G. (1992). Analysis of the association between marital relationships and health problems: an interactional perspective. *Psychol. Bull.* 112, 39–63. doi: 10.1037/0033-2909.112.1.39
- Buss, D. M. (2005). *The Handbook of Evolutionary Psychology*. New Jersey: Wiley.
- Busso, C., Lee, S., and Narayanan, S. (2009). Analysis of emotionally salient aspects of fundamental frequency for emotion detection. *IEEE T Audio Speech* 17, 582–596. doi: 10.1109/TASL.2008.2009578
- Caughlin, J. P., and Huston, T. L. (2002). A contextual analysis of the association between demand/withdraw and marital satisfaction. *Pers. Relationsh.* 9, 95–119. doi: 10.1111/1475-6811.00007
- Christensen, A. (2010). "A unified protocol for couple therapy," in *Enhancing Couples. The Shape of Couple Therapy to Come*, eds K. Hahlweg, M. Grawe-Gerber, and D. H. Baucom (Cambridge: Hogrefe), 33–46.
- Christensen, A., Eldridge, K., Cattapreta, A. B., Lim, V. R., and Santagata, R. (2006). Cross-cultural consistency of the demand-withdraw interaction pattern in couples. *J. Marriage Fam.* 68, 1029–1044. doi: 10.1111/j.1741-3737.2006.00311.x
- DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., et al. (2006). Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proc. Biol. Sci.* 273, 1355–1360. doi: 10.1098/rspb.2005.3445
- Ditzen, B., Hahlweg, K., Fehm-Wolfsdorf, G., Groth, T., and Baucom, D. H. (2011). Assisting couples to develop healthy relationships: effects of couples relationship education on cortisol. *Psychoneuroendocrinology* 36, 597–607. doi: 10.1016/j.psyneuen.2010.07.019
- Doss, B., Simpson, L. S., and Christensen, A. (2004). Why do couples seek marital therapy? *Prof. Psychol. Res.* 35, 608–614. doi: 10.1037/0735-7028.35.6.608
- Feinberg, D. R. (2008). Are human faces and voices ornaments signaling common underlying cues to mate value? *Evol. Anthropol.* 17, 112–118. doi: 10.1002/evan.20166
- Feinberg, D. R., DeBruine, L. M., Jones, B. C., Little, A. C., O'Connor, J. J. M., and Tigue, C. C. (2012). Women's self-perceived health and attractiveness predict their male vocal masculinity preferences in different directions across short- and long-term relationship

- contexts. *Behav. Ecol. Sociobiol.* 66, 413–418. doi: 10.1007/s00265-011-1287-y
- Feinberg, D. R., Jones, B. C., Law-Smith, M. J., Moore, F. R., DeBruine, L. M., Cornwell, R. E., et al. (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Horm. Behav.* 49, 215–222. doi: 10.1016/j.yhbeh.2005.07.004
- Giles, H., and Coupland, N. (1991). *Language: Context and consequences*. California: Brooks/Cole.
- Gottman, J. M. (1993). A theory of marital dissolution and stability. *J. Fam. Psychol.* 7, 57–75. doi: 10.1037/0893-3200.7.1.57
- Gottman, J. M., and Levenson, R. (1992). Marital processes predictive of later dissolution: behavior, physiology, and health. *J. Pers. Soc. Psychol.* 63, 221–233. doi: 10.1037/0022-3514.63.2.221
- Gottman, J. M., and Notarius, C. I. (2002). Marital research in the 20th century and a research agenda for the 21st century. *Fam. Proc.* 41, 159–197. doi: 10.1111/j.1545-5300.2002.41203.x
- Greenberg, L. S., and Pascual-Leone, A. (2006). Emotion in psychotherapy: a practice-friendly research review. *J. Clin. Psychol.* 62, 611–630. doi: 10.1002/jclp.20252
- Gregory, S. W., and Webster, S. (1996). A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perception. *J. Pers. Soc. Psychol.* 70, 1231–1240. doi: 10.1037/0022-3514.70.6.1231
- Hahlweg, K., Baucom, D. H., Grawe-Gerber, M., and Snyder, D. K. (2010). “Strengthening couples and families: Dissemination of interventions for the treatment and prevention of couple distress,” in *Enhancing Couples*, eds K. Hahlweg, M. Grawe-Gerber, and D. H. Baucom (Cambridge: Hogrefe), 3–30.
- Hahlweg, K., and Richter, D. (2010). Prevention of marital distress and instability: results of an 11-year follow-up study. *Behav. Res. Ther.* 48, 377–383. doi: 10.1016/j.brat.2009.12.010
- Hassan, I., McCabe, R., and Priebe, S. (2007). Professional-patient communication in the treatment of mental disorders: a review. *Commun. Med.* 4, 141–152. doi: 10.1515/CAM.2007.018
- Hawkins, A. J., Blanchard, V. L., Baldwin, S. A., and Fawcett, E. B. (2008). Does marriage and relationship education work? A meta-analytic study. *J. Consult. Clin. Psychol.* 76, 723–734. doi: 10.1037/a0012584
- Hodges-Simeon, C. R., Gaulin, S. J. C., and Puts, D. A. (2010). Different vocal parameters predict perceptions of dominance and attractiveness. *Hum. Nat.* 21, 406–427. doi: 10.1007/s12110-010-9101-5
- Hollien, H., Hollien, P. A., and de Jong, G. (1997). Effects of three parameters on speaking fundamental frequency. *J. Acoust. Soc. Am.* 102, 2984–2992. doi: 10.1121/1.420353
- Hollien, H., Green, R., and Massey, K. (1994). Longitudinal research on adolescent voice change in male. *J. Acoust. Soc. Am.* 95, 2646–2654. doi: 10.1121/1.411275
- Hollien, H., and Shipp, T. (1972). Speaking fundamental frequency and chronological age in males. *J. Speech Hear Res.* 15, 155–159.
- Hulsman, R. L., Smets, E. M. A., Karemaker, J. M., and de Haes, H. J. C. J. M. (2011). The psychophysiology of medical communication. Linking two worlds of research. *Patient Educ. Couns.* 84, 420–427. doi: 10.1016/j.pec.2011.05.008
- Irwin, J. R. (2003). Parent and non-parent perception of the multimodal infant cry. *Infancy* 4, 503–516. doi: 10.1207/S15327078IN0404-06
- Jacobson, N. S., and Christensen, A. (1998). *Acceptance and Change in Couple Therapy: A Therapist's Guide to Transforming Relationships*. New York: Norton.
- Johannes, B., Wittels, P., Enne, R., Eisinger, G., Castro, C. A., Thomas, J. L., et al. (2007). Non-linear function model of voice pitch dependency on physical and mental load. *Eur. J. Appl. Physiol.* 101, 267–276. doi: 10.1007/s00421-007-0496-6
- Junger, J., Pauly, K., Bröhr, S., Birkholz, P., Neuschaefer-Rube, C., Kohler, C., et al. (2013). Sex matters: Neural correlates of voice gender perception. *Neuroimage* 79, 275–287. doi: 10.1016/j.neuroimage.2013.04.105
- Juslin, P. N., and Laukka, P. (2003). Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol. Bull.* 129, 770–814. doi: 10.1037/0033-2909.129.5.770
- Juslin, P. N., and Scherer, K. (2005). “Vocal expression of affect,” in *The new Handbook of Methods in Nonverbal Behavioral Research*, eds J. Harrigan, R. Rosenthal, and K. R. Scherer (New York: Oxford University Press), 65–136.
- Kaiser, A., Hahlweg, K., Fehm-Wolfsdorf, G., and Groth, T. (1998). The efficacy of a compact psychoeducational group training program for married couples. *J. Consult. Clin. Psychol.* 66, 753–760. doi: 10.1037/0022-006X.66.5.753
- Kiecolt-Glaser, J. K., and Newton, T. L. (2001). Marriage and health: his and hers. *Psychol. Bull.* 127, 472–503. doi: 10.1037/0033-2909.127.4.472
- Kliem, S., Weusthoff, S., Baucom, B. R., and Hahlweg, K. (2012). “Predicting long-term risk for divorce using non-parametric conditional survival trees,” in *Poster presented at the 46th Annual Convention of the Association for Behavioral and Cognitive Therapies*, National Harbor, USA.
- Kuenzel, H. J. (2000). Effects of voice disguise on speaking fundamental frequency. *For. Linguist* 7, 1350–1771.
- Larsen, J. T., Berntson, G. G., Poehlmann, K. M., Ito, T. A., and Cacioppo, J. T. (2008). “The psychophysiology of emotion,” in *The Handbook of Emotions*, eds R. Lewis, M. Haviland-Jones, and L. F. Barrett (New York: Guilford), 180–195.
- Larson, K., and Halfon, N. (2013). Parental divorce and adult longevity. *Int. J. Public Health* 58, 89–97. doi: 10.1007/s00038-012-0373-x
- Lee, C., Baucom, B. R., Katsamanis, A., Christensen, A., Georgiou, P., and Narayanan, S. (2012). Modeling vocal interdependence during marital interaction: a vector-based approach. *Talk presented at the 46th Annual Convention of the Association for Behavioral and Cognitive Therapies*, National Harbor, USA.
- Linnman, C., Moulton, E. A., Barmetler, G., Becerra, L., and Borsoo, D. (2012). Neuroimaging of the periaqueductal gray: state of the field. *Neuroimage* 60, 505–522. doi: 10.1016/j.neuroimage.2011.11.095
- O'Brien, M. (1988). Men and fathers in therapy. *J. Fam. Ther.* 10, 109–123. doi: 10.1046/j.1988.00306.x
- O'Connor, J. J. M., Fraccaro, P. J., and Feinberg, D. R. (2012). The influence of male voice pitch on women's perceptions of relationship investment. *J. Evol. Psychol.* 10, 1–13. doi: 10.1556/JEP.10.2012.1.1
- O'Connor, J. J. M., Re, D. E., and Feinberg, D. R. (2011). Voice pitch influences perception of sexual infidelity. *Evol. Psychol.* 11, 64–78.
- Owren, M. J., and Bachorowski, J. (2007). “Measuring emotion-related arousal,” in *Handbook of Emotion Elicitation and Assessment*, eds J. A. Coan, and J. J. B. Allen (Oxford: Oxford University Press), 239–266.
- Pipitone, R. N., and Gallup, G. G. Jr. (2008). Women's voice attractiveness varies across the menstrual cycle. *Evol. Hum. Behav.* 29, 268–274. doi: 10.1016/j.evolhumbehav.2008.02.001
- Puts, D. A. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evol. Hum. Behav.* 26, 388–397. doi: 10.1016/j.evolhumbehav.2005.03.001
- Puts, D. A. (2010). Beauty and the beast: mechanisms of sexual selection in humans. *Evol. Hum. Behav.* 31, 157–175. doi: 10.1016/j.evolhumbehav.2010.02.005
- Puts, D. A., Gaulin, S. J. C., and Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evol. Hum. Behav.* 27, 283–296. doi: 10.1016/j.evolhumbehav.2005.11.003
- Puts, D. A., Jones, B. C., and DeBruine, L. M. (2012). Sexual selection on human faces and voices. *J. Sex Res.* 49, 227–243. doi: 10.1080/00224499.2012.658924
- Robles, T. E., and Kiecolt-Glaser, J. K. (2003). The physiology of marriage: pathways to health. *Physiol. Behav.* 79, 409–416. doi: 10.1016/S0031-9384(03)00160-4
- Scherer, K. R. (2003). Vocal communication of emotion: a review of research paradigm. *Speech Commun.* 40, 227–256. doi: 10.1016/S0167-6393(02)00084-5
- Scherer, K. R. (2005). What are emotions? And how can they be measured? *Soc. Sci Inform.* 44, 695–725. doi: 10.1177/0539018405058216
- Scherer, K. R. (2009). The dynamic architecture of emotion: evidence for the component process model. *Cogn. Emot.* 23, 1307–1351. doi: 10.1080/02699930902928969
- Sloan, D. M., and Kring, A. M. (2007). Measuring changes in emotion during psychotherapy: Conceptual and methodological issues. *Clin. Psychol. Sci.* 14, 307–322. doi: 10.1111/j.1468-2850.2007.00092.x
- Sokhi, D. S., Hunter, M. D., Wilkinson, I. D., and Woodruff, P. W. R. (2005). Male and female voices activate distinct regions in the male brain. *Neuroimage* 27, 572–578. doi: 10.1016/j.neuroimage.2005.04.023
- Streeter, L. A., Macdonald, N. H., Apple, W., Krauss, R. M., and Galotti, K. M. (1982). Acoustic and perceptual indicators of emotional stress. *J. Acoust. Soc. Am.* 73, 1354–1360. doi: 10.1121/1.389239
- Titze, I. (1989). Physiological and acoustic differences between male and female voices. *J. Acoust. Soc. Am.* 85, 1699–1707. doi: 10.1121/1.397959

- Tolkmitt, E., Helfrich, H., Standke, R., and Scherer, K. R. (1982). Vocal indicators of psychiatric treatment effects in depressives and schizophrenics. *J. Commun. Disord.* 15, 209–222. doi: 10.1016/0021-9924(82)90034-X
- Uchino, B. N., Cacioppo, J. T., and Kiecolt-Glaser, J. K. (1996). The relationship between social support and physiological processes: a review with emphasis on underlying mechanisms and implications for health. *Psychol. Bull.* 119, 488–531. doi: 10.1037/0033-2909.119.3.488
- Verhofstadt, L. L., Buyssee, A., Ickes, W., De Clercq, A., and Peene, O. J. (2005). Conflict and support interactions in marriage: an analysis of couples' interactive behavior and on-line cognition. *Pers. Relationsh.* 12, 23–42. doi: 10.1111/j.1350-4126.2005.00100.x
- Weusthoff, S., Baucom, B. R., and Hahlweg, K. (2013). Understanding fundamental frequency during couple conflict: an analysis of physiological, behavioral, and sex-linked information encoded in emotional arousal. *J. Fam. Psychol.* 27, 212–220. doi: 10.1037/a0031887
- Whalen, D. H. and Levitt, A. G. (1995). The universality of intrinsic f₀ of vowels. *J. Phon.* 23, 349–366. doi: 10.1016/S0095-4470(95)80165-0
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 28 March 2013; accepted: 25 June 2013; published online: 15 July 2013.
Citation: Weusthoff S, Baucom BR and Hahlweg K (2013) The Siren song of vocal fundamental frequency for romantic relationships. *Front. Psychol.* 4:439. doi: 10.3389/fpsyg.2013.00439
- This article was submitted to *Frontiers in Emotion Science*, a specialty of *Frontiers in Psychology*.
- Copyright © 2013 Weusthoff, Baucom and Hahlweg. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.