

Can we Translate Sounds into Words? A response to Leo Groarke's *Auditory Arguments: The Logic of 'Sound' Arguments*

GABRIJELA KIŠIČEK

*Faculty of Humanities and Social Sciences
University of Zagreb
Croatia
gkisicek@ffzg.hr*

Abstract: This comment to Leo Groarke's *Auditory Arguments: The Logic of 'Sound' Arguments* is a contribution to the better understanding of an auditory argument as a part of analysis of an argumentative discourse. The emphasis is on human sound i.e. prosodic features of spoken language and its argumentative function. Paper presents sort of "auditory dictionary" which might be of use in sound analysis. It also gives one possible solution of translating sound into words by using visual images as mediators.

Résumé: Mes commentaires sur l'article de Leo Groarke, *Les arguments auditifs : la logique des arguments «sonores»*, contribuent à une meilleure compréhension d'un argument auditif dans le cadre de l'analyse d'un discours argumentatif. L'accent est mis sur le son humain, à savoir, les caractéristiques prosodiques du langage parlé et sa fonction argumentative. Cet article présente une sorte de "dictionnaire auditif" qui pourrait être utile dans l'analyse sonore. Il donne également une solution possible pour traduire le son en mots en utilisant des images visuelles comme médiateurs.

Keywords: auditory argument; auditory dictionary; prosodic features; visual images

1. Introduction

In recent years there has been rising interest among the rhetoric, argumentation scholars and informal logicians in analysing the role of sound in an argumentative discourse. Sound (or auditory) argument fits in perfectly in the realm of multi-modal argumentation. First, let us answer the question: what are "modes of argumentation" and "multi-modal argumentation"? According to Groarke (2015:

p.149), ‘modes’ are defined in terms of “the ingredients used in constructing arguments.” Multi-modal discourse may thus consist of the verbal part of a message, but it can also include numerous non-verbal modes to construct an argument. So, we consider sound as one of several possible ingredients which can be a part of an argumentative discourse. However, as Groarke correctly emphasizes, this ingredient has been ignored in the past. We do not know how to deal with many different sounds that occur in an everyday argumentation; we do not know how to understand them or evaluate them. Because sounds that surround us are so inherently present and connected with the spoken language that we perceive sounds without even noticing. For instance, when we listen to the daily news, we concentrate on the verbal message and information that interests us. But suddenly we change the channel because something is bothering us, making the message difficult to understand, making a “noise” in the communication. And without exactly knowing what it is, we react (change the channel). Maybe the television presenter has a specific voice quality which enables us to concentrate on the news itself, the reading is not “logical,” has the wrong word emphasis, and many similar tiny details that television professionals know very well but are less known to an ordinary viewer. This is just an illustration of the inherent and natural connection between words and the human voice. Although, rhetoric (as Groarke writes it) has been interested in voice, pitch, intonation, etc. (Quintilian in his *Institutio Oratoriae* was the first who gave the most extensive explanation on how to use prosodic features effectively), those features have been a part of speech delivery and never seriously regarded as part of invention.

So, without much persuasion, many people will accept the notion of sound having a communicative role, some will accept the idea of sound having a rhetorical function. But, is it possible for sound to be an argument? Can we talk about auditory argument and consider it of equal importance as verbal? This is the question that Groarke is trying to answer. The starting point in which the verbal and auditory part of an argument are of equal importance is difficult to defend, and Groarke is very well aware of it, so at the end of his essay he writes:

In this essay, I have tried to show that non-verbal sounds play a significant role in reasoning, argument, and argumentation more broadly conceived. In the process I have tried to provide a series of examples that illustrate some of these roles in practice.

In my opinion, Groarke successfully argued for the importance to take sound into account when analyzing argumentative discourse. But he also pointed to some problems that prevent us from even thinking about equality between verbal and auditory.

In my response, I will try to make a step forward to this ambitious idea of equality by offering some possible solutions to some potential problems.

2. What is sound and how do we understand it?

Groarke writes:

The distinctions between “a sound,” “a noise,” “silence,” “music,” and “normal and abnormal” sounds are, therefore, drawn in different ways at different times and places.

Out of context, a sentence such as this might give sceptics a reason to say, “Yes, this ambiguity shows that sounds cannot be interpreted correctly in an argumentative discourse because different people will the same sound understand differently.” And one can answer, “Yes but words can also be ambiguous.” However, if the distinction between normal and abnormal is so individually, or culturally, or temporally interpreted, then it makes it very hard to analyse or evaluate sound arguments. In this section, I will concentrate on the achievements made in the scientific investigation of how we might objectively interpret non-verbal human sounds. Numerous studies and empirical research have provided us a sort of an “auditory dictionary.” If we hear a certain voice type, we connect it to a certain personality type. It is, of course, stereotypical perception. But it has been replicated and repeatedly shown that some voices are going to correlate with a specific personality perception.

One of the pioneer studies was by Heidenberg in 1964 (according to Hickson et al., 2004), who theorized 11 voice types and their stereotypes. Of the 11 voice types only one was perceived as “good,”

and it is the one associated with proper breathing, articulation, tongue position, control of pitch and resonance.

The other 10 voice types are:

VOCAL TYPE	MALE	FEMALE
Breathy	artistic, younger, feminine	Sexy
Tense	anxious, nervous, uncooperative, less intelligent, high-strung	anxious, nervous, uncooperative, less intelligent, high-strung
Breathy-tense	weak, nervous	weak, nervous
Nasal	whiny, argumentative, lazy	whiny, argumentative, lazy
Denasal	stuffy, boring, speaker with a cold	stuffy, boring, speaker with a cold
Orotund	energetic, pompous, humor-less, proud, authoritative	energetic, pompous, humor-less, proud, authoritative
Flat	bored, withdrawn, sluggish	bored, withdrawn, sluggish
Thin	immature, sensitive, emotional	especially for women: immature, sensitive, emotional
Throaty	sophisticated, less intelligent, careless, older	sophisticated, less intelligent, careless, older
Fronted	artificial, aloof	artificial, aloof

Berry (1991, 1992) did empirical research on specific voice types and provided us with similar results:

TYPE	MALE	FEMALE
attractive voice	deep – associated with strength, assertiveness, dominance	not necessarily deep but with balanced vocal qualities – warm, kind, honest
babyish voice	less competent, less powerful with no leadership qualities but more honest and warm	less competent, less powerful with no leadership qualities but more honest and warm

Further on, many authors dealing with nonverbal communication conducted research that added to the auditory dictionary.

PERSONALITY TRAITS	PROSODIC QUALITIES
Extroverted person (Lippa 1998, Seigman, 1987)	fluent speech, shorter pauses, fewer hesitations, faster, louder, wider pitch range, more dynamic
Introverted person (Lippa 1998, Seigman, 1987)	opposite from extroverted person
Masculine person (associated with a tendency to certain professions, hobbies, activities) – Lippa 1998	deeper voice, slower tempo, louder speech, less expressive (for men); only deeper voice for women
Aggressive person (Friedman, Hall & Harris 1985)	faster speech, frequent changes in speech rate, frequent pauses, staccato rhythm, explosive articulation
Dominant person (Harrigan, Gramata, Luck and Margolis, 1989; Tusing and Dillard 2000)	faster and louder speech, deeper voice
Physically attractive person, more sexually active person with more sexual partners	attractive voices: balanced spectrum, deeper voice

<p>Sexy person (Oguchi and Kikuchi 1997)</p>	<p>Sexy voice – low pitched voice with a small pitch range Women perceive high-pitched male voices to be unattractive and prefer deeper male voices (Collins & Feeney 2000). The preference for lower-pitched male voices also appears to be replicated cross-culturally and in Non-Western societies (Apicella, 2007). Men, on the other hand, perceive higher-pitched female voices as sounding more attractive (Collins and Missing 2003)</p>
--	---

Also, there has been a lot of progress in proving and providing empirical evidence for the connection between specific prosodic features of human voice and perception of emotions. Evidence of cross-cultural agreement in how basic emotions are recognized from a speaker’s vocal expressions has been reported (Davitz, 1964; Van Bezooijen et al. 1984; Bachorowski, 1999; Scherer, 2003; Thompson and Balkwill, 2006; Scherer et al. 2011). Scherer et al. (2011) presented 30 emotionally-inflected but semantically-anomalous “pseudo-utterances” produced by four German actors to native speakers of nine different languages. The authors found that all listener groups recognized fear, joy, sadness, anger and “neutral” utterances strictly from prosody at above chance accuracy levels. So, it has been proven that when adults listen to a foreign language they can successfully infer the speaker’s emotional state strictly from their vocal inflections while speaking, consistent with previous findings (Davitz, 1964; Van Bezooijen et al. 1984; Johnson et al, 1986; Banse & Scherer, 1996; Scherer et al. 2011, Laukka et al 2013)

EMOTION/CONDITION	PROSODIC QUALITIES
anxiety (combination of fear, uncertainty, distress, apprehension, and worry)	dysfluencies of speech (e.g., increased pausing), lowered pitch variability, higher pitch and tense voice, increased jitter (i.e., pitch perturbations) Eldred and Price 1958; Hofmann et al. 1997; Kasl and Mahl 1965; Lewin et al. 1996; Mahl 1956; Pope et al. 1970, Fuller et al. (1992)
confident	louder and faster (Kimble & Seidel 1991)
deceptive voice	higher frequency (Ekman et al. 1976; Anoli, Ciceri 2002)
joy	wider pitch range, higher pitch, tense voice (Novak& Vokral, 1993)
anger	tense voice, louder (Novak& Vokral, 1993)
sadness	looseness of larynx, slower, less intensity (Novak& Vokral, 1993), low fundamental frequency (Laver, 1980)
withdrawn, distant	hoarse voice quality (Tanner and Tanner, 2004)

Keeping in mind that all of these results are well described, well known and available, we can be sure that they are going to be (and are) used in the public sphere. Groarke implies as much, but in the rhetorical context when creating a certain disposition for the audience (*pathos*) and gives an example of a potential conversation between spouses.

Example 26: My spouse is more prone to agree with me when she is in a good mood. She loves classical music, so I play her a new recording of the first movement of Beethoven's 7th symphony before I argue that we should buy a new automobile.

In this case, a musical performance sets the stage for argument. It is notable because it is not a neutral background but is used in an attempt to influence the audience to whom my arguments are directed. In many other cases of public argument – at political rallies, in advertisements, in documentary film – music or other sounds are used in a similar way.

However, knowledge from the extensive experiments in nonverbal communication can be successfully used not just as a part of *pathos*, but also as a part of *ethos* and *logos*.

So, prosodic features which, as mentioned before, contribute to the perception of a speaker's character may influence the perception of her *ethos*. For instance, if a politician claims that he/she would be the best choice for the next president because he/she is strong, determined and capable of dealing with all the problems country is facing, we can imagine that, for example very nasal voice will contradict his claim on strength and capability he/she possesses. Nasality makes for a vocal characteristic considered to be particularly undesirable in public speaking. As Bloom, Zajac & Titus (1999, p. 279) state:

Highly nasal voices were rated as being lower in "status" (occupation, ambitious, intelligent, educated, influential), lower in social solidarity (friendly, sympathetic, likeable, trustworthy, helpful), and were negatively correlated with perceptions of persuasiveness.

It is not meant that a speaker with nasal voice has no possibility of ever becoming a president, but merely that strategy of building *ethos* should be based on more than pure claim of personal strength and confidence. Personality traits of politicians are an important part of their *ethos* and are commonly reflected in the prosodic features of her speech.

Of course, the most important notion of the contributions made by nonverbal communication research is the part of a *logos*. How sound can become the essential part of an argument reconstruction, i.e. becoming an auditory argument, as Groarke so accurately defines it:

Auditory (acoustic or sonic) argument is an attempt to provide rational evidence for a conclusion using non-verbal sounds instead of or (more frequently) in addition to words.

Along with this definition, Groarke gives several very good examples corroborating the claim that auditory arguments do exist. We will return to some of these examples in the next section.

To conclude the problem of sound ambiguity: Of course, it is possible that sound is ambiguous and can be interpreted differently by different people from different cultures. But having so much data, so many experiments showing us above average agreement between different people in the perception of a human voice, gives us much more hope in dealing with human sound as a reliable ingredient of an argumentative discourse.

3. Argument *about* sound vs. auditory argument

Examples 3 and 4 in Groarke's paper illustrate the difference between argument *about* sound and auditory argument. Discussing the quality of an orchestra and claiming that they played a superb performance by corroborating it with comparisons, arguments from authority and so on, is different than providing auditory evidence for the claim that someone sang superbly.

In this section, I want to add the rise and success of a specific science called forensic phonetics (International Association for Forensic Phonetics and Acoustics, www.iapfa.com). It is worth mentioning this relatively new branch of science because they use auditory arguments to corroborate a claim that a specific person is (or is not) a perpetrator of a crime. The process of determining the perpetrator on the basis of comparison between two recordings is far more complicated than in popular television crime series. However, this simplified version of forensic phonetics will suit us as an example of how auditory arguments are used not just in everyday argumentation but also in a judicial context.

Example 1: Person x is the one who made a telephone threat. We have the recording of a telephone call, and we have a recording of the police interview. Listen to these two recordings, compare it, and you will hear that this is the same person.

This is a simplified version of how a legal argument may proceed, however, representative of argumentation in a court of law. To corroborate the claim that the person *x* is the same voice on two recordings is a complicated and difficult task which involves experts and many acoustic testing and in the end is never 100 percent accurate. But, after all the measurements in a lab, all the testing the argument is going to be constructed in the following way:

Person x quality because his voice is very likely the voice of a person who made the call.

4. Analysing and assessing auditory arguments

One of the most difficult problems we need to overcome in dealing with auditory arguments is to find an appropriate way of analysing and assessing them.

Groarke writes:

We could analyse and assess auditory arguments by looking for a way to turn them into verbal arguments which can then be analysed and assessed as verbal arguments. The most obvious way to do so is by describing the sounds that they contain: a process that turns these sounds into words (one might compare attempts to deal with visual arguments by turning them into verbal arguments that describe the visuals in question). This approach is problematic because non-verbal sounds are (like visuals) notoriously difficult to express in words, there being no exact way to translate them into words.

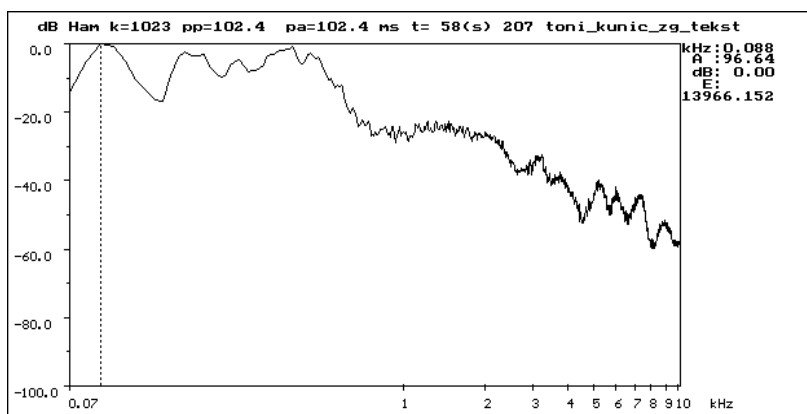
Groarke gives two possibilities:

- 1) One way to produce exacting verbal equivalents of auditory arguments like these is by relying on ostension, using language as a verbal pointer that identifies the non-verbal sounds in question (ornithologists example)
- 2) The other way to verbalize auditory arguments is by replacing their non-verbal sounds with descriptions of these sounds. In doing so, one does translate sounds into words, but in most cases, this comes at a great cost, for descriptions usually fail to precisely and completely convey the sound in question.

I would propose a third possibility in which, unlike the second, we do not lose content with translation into words, but on the contrary,

we might even gain a clearer understanding of the content. In some cases, it is possible to translate sound into words by using visual images as mediators.

Let us consider again an example of a telephone threat. The prosecution claims that x is guilty of making a telephone threat and they support the claim with an argument from authority, i.e. forensic phonetician. However, an expert must make the argument. It is not enough just to say, “I have listened to both recordings (the telephone and the police interview), compared it, and I claim it is the same person.” The argument must be stronger. The claim must be further elaborated, and non-verbal sounds must be translated into words. To make the transition from sound to verbal easier, images might be used. At the beginning of the forensic phonetics, some believed that “voiceprints” exist and that they are equal to fingerprints. Although this is not true due to many possible “intraspeaker” variations, images of a voice do exist and are called spectrums.



They can be of help when arguing that the voice of person x is the same on both recordings. This is just one example of making a sound visual before verbal. In addition to voice quality (showed on the image above), it can be used to show intonation patterns, pitch range and all of the prosodic features discussed.

5. Analysing auditory arguments

In many cases when we have no problem in recognizing the sound (sound of an animal, car engine, alarms or sirens) we can still have a problem assessing the validity of an argument made by sound. Groarke writes:

A good auditory argument (as well as verbal) must be well formed.

- When arguments are well formed we need to answer the question of whether they are good arguments by considering whether they satisfy the normal criteria for good argument.
- by looking for standard fallacies (straw man, post hoc, ad hominem, etc.); by asking the critical questions associated with argument schemes (argument by analogy, causal reasoning, etc.); by applying standard accounts of deductive and inductive validity; and so on.

In this stage of dealing with auditory arguments, the best idea is to borrow assessment tools from verbal argumentation. It is possible to understand auditory arguments as a part of different argumentation schemes. One of the examples that Groarke uses at the beginning of his paper is an argument from sign. A medical doctor listens to a patient's heart and concludes that she has a heart condition. So, to assess the validity of that argument, we can use critical questions like:

Are there other events that would more reliably account for the sign? What is the strength of the correlation of the sign with event signified? (Walton 2006)

And we can use standard patterns to (re)construct a specific argument scheme, like the one used for argument from sign:

SPECIFIC PREMISE: A (a finding) is true in this situation

GENERAL PREMISE: B is generally indicated as true when its sign, A is true.

CONCLUSION: B is true in this situation

As Groarke puts it:

The question of whether there are auditory variants of all the criteria used to judge the validity or invalidity of verbal arguments merits

further study. So too does the question whether there are any unique schemes of argument that are intrinsically auditory.

I can conclude with similar thoughts. We are on the way of drawing attention to the importance of sound in argumentation, but many questions still need to be answered. Also, this topic would benefit from empirical research which will help us answer at least some of the many questions that remain. In the end, Groarke's essay starts a path for a better understanding of an auditory ingredient in multimodal argumentative discourse and provides good reason to continue with the same enthusiasm.

References

- Anolli, L., & Ciceri, R. 2002. Analysis of the vocal profiles of male seduction: From exhibition to self-disclosure. *Journal of General Psychology*, 129: 149–169.
- Apicella, C. T. 2007. Voice Pitch Predicts Reproductive Success in Male Hunter-Gatherers. *Biology Letters* 3: 682-684.
- Bachorowski, J.-A. 1999. Vocal expression and perception of emotion. *Current Directions in Psychological Science*, 8: 53–57.
- Banse, R., & Scherer, K. R. 1996. Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3): 614–636.
- Berry, D. S. 1991. Accuracy in social perception: Contributions on facial and vocal information. *Journal of Personality and Social Psychology*, 62: 298-307.
- Berry, D. S. 1992. Vocal types and stereotypes: Joint effects of vocal attractiveness and vocal maturity on person perception. *Journal of Nonverbal Behavior*, 18: 187-197.
- Bloom, K., Zajac, D.J. & Titus, J. 1999. The influence of nasality of voice on sex-stereotyped perceptions. *Journal of Nonverbal Behavior* 23(4): 271-281.

- Collins, S. A., Missing, C. 2003. Vocal and Visual Attractiveness are Related in Woman. *Animal Behavior* 65: 997-1004.
- Collins, N. L., & Feeney, B. C. 2000. A safe haven: An attachment theory perspective on support seeking and caregiving in intimate relationships. *Journal of Personality and Social Psychology*, 78: 1053–1073.
- Davitz, J. R. 1964. Auditory correlates of vocal expression of emotional feeling. In J. R. Davitz (Ed.), *The communication of emotional meaning* (pp. 101–112). New York: McGraw-Hill.
- Eldred, S. H., & Price, D. B. 1958. A linguistic evaluation of feeling states in psychotherapy. *Psychiatry*, 21: 115–121.
- Ekman, P., Friesen, W.V., & Scherer, K.R. 1976. Body movement and pitch in deceptive interaction. *Semiotica*, 16: 23–37.
- Friedman, H. S., Hall, J. A., & Harris, M. J. 1985. Type A behavior, nonverbal expressive style, and health. *Journal of Personality and Social Psychology*, 48(5): 1299-1315.
- Fuller, B. F., Horii, Y., & Conner, D. A. 1992. Validity and reliability of nonverbal voice measures as indicators of stressor-provoked anxiety. *Research in Nursing and Health*, 15: 379–389.
- Groarke, L. 2015. Going Multimodal: What is a Mode of Arguing and Why Does it Matter? *Argumentation*, 29: 133-155.
- Harrigan, J. A., Grammata, J. F, Luck, K. S., & Margolis, C. 1989. It's how you say it! Physicians Vocal Behavior. *Social Science & Medicine*, 28: 87-92.
- Hickson, M., Stacks, D. & Moore, N. 2004. *Nonverbal communication – Studies and Applications*. Los Angeles: Roxbury Publishing Company.
- Hofmann, S. G., Gerlach, A. L., Wender, A., & Roth, W. T. 1997. Speech disturbances and gaze behaviour during public speaking in subtypes of social phobia. *Journal of Anxiety Disorders*, 11: 573–585.
- Johnson, W. F., Emde, R. N., Scherer, K. R., & Klinnert, M. D. 1986. Recognition of emotion from vocal cues. *Archives of General Psychiatry*, 43: 280–283.

- Kasl, S. V., & Mahl, G. F. 1965. The relationship of disturbances and hesitations in spontaneous speech to anxiety. *Journal of Personality and Social Psychology, 1*: 425–433.
- Kimble, C. E., & Seidel, S. D. 1991. Vocal signs of confidence. *Journal of Nonverbal Behavior, 15*: 99–105.
- Knapp, M. L., Hall, J. A. & Horgan, T. G. 2013. *Nonverbal communication in Human Interaction*. Boston: Wadsworth Publishing Company.
- Laukka, P., Elfenbein, H. A., Söder, N., Nordström, H., Althoff, J., Chui, W., et al. 2013. Cross-cultural decoding of positive and negative non-linguistic vocalizations. *Frontiers in Psychology, 4*: 353.
- Laver, J. 1980. *The phonetic description of voice quality*. Cambridge, England: Cambridge University Press.
- Lewin, M. R., McNeil, D. W., & Lipson, J. M. 1996. Enduring without avoiding: Pauses and verbal dysfluencies in public speaking fear. *Journal of Psychopathology and Behavioral Assessment, 18*: 387–402.
- Lippa, R. 1998. The nonverbal display and judgment of extraversion, masculinity, femininity, and gender diagnosticity: A lens model analysis. *Journal of Research in Personality, 32*(1): 80-107.
- Litman, D., Pope, B., Blass, T., Siegman, A. W., & Rahe, J. 1970. Anxiety and depression in speech. *Journal of Consulting and Clinical Psychology, 35*: 128–133.
- Mahl, G. F. (1956). Disturbances and silences in the patient's speech in psychotherapy. *Journal of Abnormal and Social Psychology, 53*: 1–15.
- Novak, A. & Voklar, J. 1993. Emotions in the Sight of Long-Time Averaged Spectrum and Three-Dimensional Analysis of Periodicity. *Folia Phoniatica et Logopaedica, 45*: 198–203.
- Oguchi, T., Kikuchi, H. 1997. Voice and interpersonal attraction. *Japanese Psychological Research, 39*(1): 56-61.
- Scherer, K. R. 2003. Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*: 227–256.

- Scherer, K. R., Clark-Polner, E., & Mortillaro, M. 2011. In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. *International Journal of Psychology*, 46(6): 401–435.
- Siegmán 1987. The telltale voice: Nonverbal messages of verbal communication. In A. W. Siegmán, & F. Stanley (Eds.), *Nonverbal behavior and communication* (2nd Ed.) (pp. 351-433). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Thompson, W., & Balkwill, L.-L. 2006. Decoding speech prosody in five languages. *Semiotica*, 158(1/4): 407–424.
- 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.
- Van Bezooijen, R. A. 1984. *Characteristics and recognizability of vocal expressions of emotion*. Doctoral dissertation. Dordrecht, The Netherlands: Foris Publications.
- Walton, D. 2004. *Fundamentals of Critical Argumentation*. Cambridge: Cambridge University Press.