INTRODUCTION TO SPECIAL ISSUE

Challenges in studying prosody and its pragmatic functions: Introduction to *JIPA* special issue

Oliver Niebuhr
SDU Electrical Engineering, Mads Clausen Institute, University of Southern Denmark
oln@sdu.dk

Nigel G. Ward
Department of Computer Science, University of Texas at El Paso
nigel@utep.edu

The impetus for this special issue was an all-day event at the 2015 meeting of the International Pragmatics Association: The Panel on Prosodic Constructions in Dialog. This event had several motivations: (i) we have enormous data sets and tools to process them, but as a field we lack clear roadmaps for how to exploit these sets and tools to improve our understanding; (ii) we know that prosody is more than just the single stream of intonation, but we find it hard to accurately describe multistream phenomena; (iii) we have observed how prosody serves many dialog and interactional functions, but cannot yet really model how; and (iv) we have various schools of thought, each wielding its own methods, but we have difficulty reconciling and connecting their various insights.

This diversity of approaches is a strength, but also a weakness, to the extent that it impedes communication and progress. In this special issue we bring together four contributions, from very different perspectives, spanning the ways researchers frame problems in the prosody of dialog. We have worked with the authors to make their research accessible to researchers from different traditions. Thus, this special issue is for all who wish to broaden their perspective on the prosody of dialog. This includes of course those already involved in prosody research or phonetics more generally but also others.

In particular, scientists from other fields see the importance of understanding prosody. It is fundamental to face-to-face dialog, which is a prototype of human social interaction, and rich in examples of the interpersonal coordination skills that make humans what we are, different from all other animals (Clark 1996, Levinson 2006, Dale et al. 2013, Melis et al. 2016). These abilities are of importance for psycholinguistics, sociology and the other sciences, and are of more than academic interest. There are many active researchers and practitioners working in micro-analysis and conversation analysis, including for such practical purposes as couples therapy, optimizing service interactions, improving workplace communication,

doi:10.1017/S0025100318000014

and diagnosing, assessing, and treating communication disorders. There is also a great public demand for knowledge of how to communicate more effectively, with many practical questions as yet unanswered.

Another community that needs models of prosody is the speech engineering community. For example, speech synthesis is now highly intelligible, but it remains so rigid that its use in dialog applications makes them awkward and tiring for users. Prosody is recognized as a key problem here (Collier 2015). Can linguistic models contribute to solving this problem? Not everyone thinks so: there has been a recent burst of progress in black-box models of prosody for synthesis (Zen, Senior & Schuster 2013). These models are nothing more than generic neural-network architectures applied to huge corpora, from which they learn thousands of parameters that together enable the generation of fairly natural-sounding prosody, at least for read-style speech. Nowhere in these models is there any recognizable representation of knowledge about prosody. The same is true for many other practical tasks – recognizing emotion from prosody, recognizing turn-taking patterns, diagnosing medical conditions involving prosodic differences and deficits, and so on. In general, when the inputs and outputs are clear, knowledge-free machine-learning algorithms outperform sophisticated knowledge-rich models. However, black-box models do not meet all needs.

In particular, speech communication is about more than just the transfer of information: there are vital social and pragmatic functions that today's dialog applications are completely insensitive to. Ironically, it is the growing number of interactive machines in our everyday life that make this issue all the more salient. Speech synthesis and recognition systems already enable us to interact with machines at a basic level: we can ask questions and get answers or vice versa. But do we really LIKE talking to machines? Surely not; and one of the main reasons is that machines are still unable to produce and perceive social and/or pragmatic forms and functions of prosody (Mayo, Clark & King 2011, Wolters et al. 2014). While speech scientists are well aware of this deficit, industry is also becoming increasingly concerned with these issues, including issues of voice design and dialog design, across automotive, medical, and many other applications areas (Chebat et al. 2007, Wolters et al. 2014, Sandry 2015, Fischer 2016, Niebuhr, Tegtmeier & Brem 2017, Rodero 2017). While some of these issues may in future succumb to black-box models, there are other applications – like helping adult learners to be effective in a new language, and producing highly-responsive dialog agents (Ward & DeVault 2016) – where we absolutely need proper, knowledge-rich models of prosody. In short, there is a need to improve our understanding and modeling of prosody, especially with respect to the prosody of dialog.

A challenge for the authors of this special issue was to consider prosody beyond just intonation. As discussed below, they addressed this to different degrees and in different ways, but all of their observations provide material for, and indeed challenges for, attempts to model prosody. The second major challenge for them was to work towards analyses in which the social and pragmatic functions are central. Below, after a brief historical survey that discusses why prosody is so hard to model accurately, we revisit these points, discussing fundamental challenges highlighted by the papers in this special issue and describing our hopes for future work.

1 Some perspective on prosody and pragmatic functions

To set the stage for the papers and to explain why we stress the role of pragmatic function as a prerequisite to prosodic modeling, this section takes a brief look at some historical developments in intonation research. Figure 1 provides an example-based overview of intonational representations. Starting with the strategy that Tuscan monks used in the 15th century for remembering the Gregorian chants associated with Bible texts (in this way inventing the modern concept of musical notes, Kelly 2014), we see the styles used in the

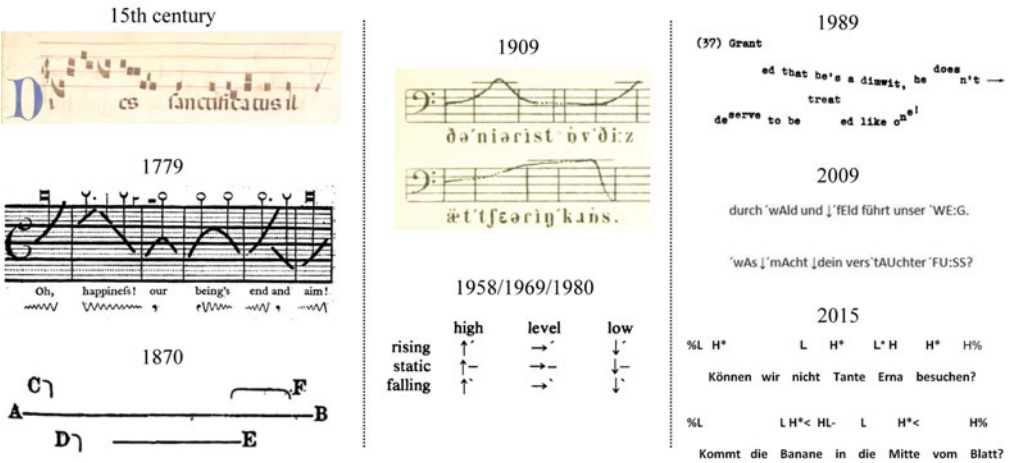

Figure 1 (Colour online) Representations of intonation through history.

early works of Steele (1779), Caswell (1870), Jones (1909), the meandering text of Bolinger (1989), the tonetic stress marks of the British School of intonation (Kingdon 1958), GAT2 (Selting et al. 2009) and the DIMA variant of the autosegmental–metrical framework (Ladd 2008, Kügler et al. 2015). This is just a sampling of the enormous range of methods that have been invented and used to represent and examine the melody of speech; and of course today we still have an enormous range of competing representation concepts and models.

This diversity of intonational representations contrasts sharply with the situation for the segmental aspects of speech. This is true even though all of speech is, ultimately, composed of 'changes in the cavities of vocal tract – openings or closings, widenings or narrowings, lengthenings or shortenings' (Liberman & Whalen 2000: 188). Nevertheless, description in terms of individual sound segments is far more convenient, and has remained as a heuristic instrument of international scholarly consensus, as in the International Phonetic Alphabet (IPA).

A key question is: Where does the relative stability and consensus in the representation of speech segments come from, and what does its absence for prosody imply for research? While many factors are doubtless involved, including the power of an established orthography and the tactile feedback of apical articulations and various oral-cavity constrictions that are lacking for prosody, we think the most important source of stability is the unit of the lexical item (Bolinger 1963). The existence of words enables articulatory sound patterns to ground out in solid semantics. That is, they are connected to and grounded in concepts understandable and transparent to all native speakers of a language, even though these concepts are often abstract or multifaceted (Harnad 1990, Lakoff 1990). For example, with semantics in the background it is relatively easy to determine that Dutch ['ɛɪxələk], ['ɛɪxk] and ['ɛɪk] are all variants of the same 'thing' (*eigenlijk* 'actually', Ernestus & Smith, 2018; see also Hawkins 2003). With such semantic grounding, syntagmatic boundaries within sound patterns can be defined consistently and (allowing for some coarticulatory tolerance) mostly unambiguously. Moreover, the semantic grounding serves as a constant point of reference to identify, describe, and model variation.

In comparison, in prosody, we constantly struggle with the question of whether some relatively small differences should be interpreted as belonging or not belonging to the same 'thing', that is, whether some variation is phonetic or phonological in nature. For example, autosegmental–metrical intonation approaches of German distinguish four pitch-accent categories in the Oldenburg model, five pitch-accent categories in the Stuttgart model,

and six categories in the GToBI model (Mayer 1995, Grice & Baumann 2002, Peters 2014); see also the discussions in Kohler (2005) and Rathcke & Harrington (2010). One reason for this is that we lack points of reference as good as those we have for lexical items; and one reason for this, in turn, is that intonational meanings are typically of a pragmatic nature and as such less tangible and definable than word meanings. However, the pragmatic meanings of intonational patterns have been a subject of intense and controversial discussions (Gussenhoven 2004, Ladd 2008, Arvaniti 2011, Prieto 2015), and it is to be hoped that better-defined pragmatic meanings will complement and eventually replace the roles that behavioral data – like reaction times and discrimination abilities – currently play in defining within- and between-category variation for prosodic phonology.

Therefore, as in the domain of morphosyntax, intonation research needs much closer integration of analyses of forms AND analyses of functions (Arvaniti 2016), so that ultimately the latter can provide solid grounding, as in lexical semantics. If we think of forms and functions as two sides of the same coin, in prosody, our understanding and modeling of one side of the coin has a decisive influence on understanding and modeling the other side of the coin. To date the success of black-box models has been in domains where this complexity can be ignored.

Against this background, the present special issue makes significant progress by providing new insights into both phonological representation AND pragmatic function.

2 Issues and challenges

First of all, we would like to emphasize that, in addition to their theoretical import, the papers in this special issue contribute a wide range of empirical facts to our understanding of prosody. The pragmatic functions considered include obviousness, greeting, reprimanding, self-repair, upgraded assessments and sarcasm, among others; the languages studied are Peninsular Spanish, Colombian Spanish, French, and British English. Across the differences in topic and approach, all four papers contribute to the goals of this special issue: all identify meanings for prosodic forms, all move beyond the traditional concerns of prosody research to explore dialog-related, interactional, and social meanings, and all examine prosody beyond just intonation. We would like to highlight five overarching issues.

One major issue across the papers relates to the stability of prosodic forms. It would be convenient if meaningful prosodic forms were always realized in the same way, but of course, as already noted, things are not this simple. Classic examples are the issue of truncation and compression of phrase-final f0 movements or, more generally, the way intonational forms can stretch to cover the necessary time span or adjust to align with the phonetic structures of stressed syllables (Arvaniti, Ladd & Mennen 1998, Wichmann, House & Rietveld 2000, Atterer & Ladd 2004, Rathcke 2016). In this special issue, Francisco Torreira and Martine Grice's paper illustrates how complex this can be: they describe a meaningful melody that may, depending on utterance length, be partly truncated, and that contains a tonal component that may, depending on the lexical content, surface as pitch accent or as a boundary tone.

Aoju Chen and Lou Boves' paper is also relevant to the question of stability of form. The authors find that the intonational expressions of sarcasm vary with the syntactic form of the vehicle utterance. For example, sarcasm seems to be expressed with low pitch on a focused word in tag questions but not in declaratives. Along similar lines, Clara Huttenlauch, Ingo Feldhausen and Bettina Braun find that, for two pragmatic functions, greeting and seeking confirmation, the prosodic forms observed vary with the lexical content used. Thus the same function was realized differently when using a pure greeting, *hola* 'hello', versus when calling by name, for example *Manolo*. Not only do Huttenlauch et al. find differences in which intonation contours are most frequently used, they also find differences in spectral tilt

(as a measure of voice quality) and pitch range. They argue that these facts are incompatible with models that assume a simple, direct form–function mapping in prosody.

The second issue is that of the role of context. While it would be convenient if prosodic forms had consistent meanings across contexts, variations in nuance and implications are known to be common (see Bolinger 1989), which brings many well-known challenges. But Rasmus Persson's work on the French *'accent d' insistance'* goes much further: he presents a meaningful prosodic form whose meaning appears to shift entirely depending on the discourse context: varying from conveying simple indication of receipt of new information to intensification and to self-correction. He reaches a radical conclusion: that intonational units have no inherent meanings at all, instead serving merely as a kind of connecting elements or hubs between other semiotic resources, including the sequential context and the lexical items chosen. If this is true – that is, if the common practice of ascribing context-independent meanings to prosodic forms is inadmissible – our field faces huge challenge, not least because of the infinite diversity of contexts and pragmatic goals. Nevertheless, we see hope. While a complete theory of meaning is a boundless task, the aspects of meaning relevant to prosody are more limited, and are heavily skewed to social and interactional meanings. With increasing attention to these aspects of language, we hope to see continuing progress towards effective modeling of context.

The third issue we would like to highlight is that of methods. Readers will find in these papers a broad sampling of methods and approaches to prosody, reflecting major differences in evidence gathering, reasoning, and conclusions. Each of our authors has adapted, combined, and extended standard methods to better suit their aims, but each is explicit about the fact that there are still limitations with the techniques used. At the same time, it is clear that all have an equally-strong awareness of the limitations of rival methods. The problem comes when trying to integrate the insights arising from different approaches. Today this is a daunting task: Persson (Section 6) goes so far as to suggest that 'interactional linguistics and intonational phonology may be dealing with different empirical realities' because of methodological differences. But as scholars we should not accept this: our challenge is to close the gaps, and discover the underlying truth, that today we are only glimpsing from a few different angles. We hope readers will be inspired to develop new, integrative approaches to the study of prosody, or otherwise work to help put the various findings into one overarching big picture.

The fourth issue is that of the relation between intonation and the rest of prosody. Historically pitch has been given priority, with other prosodic features receiving less attention. Recently a radically opposite view has become common. For example, in the prosodic-constructions approach to description (Hedberg, Sosa & Fadden 2004; Ogden 2010, 2012; Day-O'Connell 2013; Niebuhr 2013b, 2015; Rao 2013; Ward & Gallardo 2017), pitch features have no special status; instead they are one facet among others. Some recent models of aspects of prosodic perception focus on how various acoustic parameters may merge into a much smaller number of perceptual qualities, and how parameter trade-offs within this merger are organized. Examples include the intensity-weighted tonal center of gravity of Barnes et al. (2012), and the Contrast Theory of Niebuhr (2013a). It is also the case that in speech technology essentially all applications that involve prosody use multi-stream modeling. For example, automatic recognition systems of social (pragmatic) signals in speech invariably obtain the 'highest prediction accuracies ... by combining many features' (Litman & Forbes-Riley 2006: 586). Regarding this question, of the relation between intonation and the rest of prosody, the papers in this special issue take a middle ground. They all start with observations about intonation, but they also consider how other prosodic elements relate and contribute. Looking ahead, finding ways to integrate insights and models derived from intonation-centric approaches with those from comprehensive approaches will be a challenge.

The fifth issue is that of overcoming or leveraging the essentially bidirectional nature of inquiry in this area. Going back to the two-sided coin analogy, Torreira and Grice start with the pragmatic side, with the function of signaling obviousness, and by diligent examination of the associated forms, discover a pattern of variation that they might never have seen if

they had focused entirely on the side of forms. Persson starts from the other side of the coin, focusing on one form, namely salient-initial/low-primary accent, and by diligent examination of the associated functions, in actual dialog, also discovers something unexpected. We foresee an era of rapid evolution in the study of the prosody, as more researchers more thoroughly examine both sides of the coin.

3 Outlook

This decade is an exciting time for prosody: we have new tools, new methods, and new needs, but also some of the same knotty problems in modeling.

In five-to-ten years, we expect that things will look very different. We will have models which are both theoretically satisfying and practically useful, both suitable for big-data analyses and accurate for describing specific productions, both easy to understand and empirically testable. It will not be easy getting there, not least because of the challenges raised by the papers in this special issue, which we hope will inspire and incite discussion and progress, but it will be a lot of fun.

Acknowledgments

We thank our authors, our reviewers, Richard Ogden and the *JIPA* Editor, Amalia Arvaniti, for their help with this special issue.

References

Arvaniti, Amalia. 2011. The representation of intonation. In Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 757–780. Chichester: John Wiley & Sons.

Arvaniti, Amalia. 2016. Analytical decisions in intonation research and the role of representations: lessons from Romani. *Laboratory Phonology* 7, 1–43.

Arvaniti, Amalia, D. Robert Ladd & Ineke Mennen. 1998. Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of Phonetics* 26, 3–25.

Atterer, Michaela & D. Robert Ladd. 2004. On the phonetics and phonology of "segmental anchoring" of f0: Evidence from German. *Journal of Phonetics* 32, 177–197.

Barnes, Jonathan, Nanette Veilleux, Alegjna Brugos & Stefanie Shattuck-Hufnagel. 2012. Tonal Center of Gravity: A global approach to tonal implementation in a level-based intonational phonology. *Laboratory Phonology* 3, 337–383.

Bolinger, Dwight L. 1963. The uniqueness of the word. *Lingua* 12, 113–136.

Bolinger, Dwight L. 1989. *Intonation and its uses: Melody in grammar and discourse*. Stanford, CA: Stanford University Press.

Caswell, Jesse. 1870. Treatise on the tones of the Siamese language. *Siam Repository* 2, 93–101.

Chebat, Jean-Charles, Kamel El Hedhli, Claire Gélinas-Chebat & Robert Boivin. 2007. Voice and persuasion in a banking telemarketing context. *Perceptual and Motor Skills* 104, 419–437.

Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.

Collier, Rene. 2015. Prosodic analysis: A dual track? In Jan van Santen, Richard Sproat, Joseph Olive & Julia Hirschberg (eds.), *Progress in speech synthesis*, 325–329. Berlin: Springer.

Dale, Rick, Riccardo Fusaroli, Nicholas Duran & Daniel C. Richardson. 2013. The self-organization of human interaction. *Psychology of Learning and Motivation* 59, 43–95.

Day-O'Connell, Jeremy. 2013. Speech, song, and the minor third. *Music Perception* 30, 441–462.

Ernestus, Mirjam & Rachel Smith. 2018. Qualitative and quantitative aspects of phonetic variation in Dutch *eigenlijk*. In Francesco Cangemi, Meghan Clayards, Oliver Niebuhr, Barbara Schuppler

& Margaret Zellers (eds.), *Rethinking reduction: Interdisciplinary perspectives on conditions, mechanisms, and domains for phonetic variation* (Phonology and Phonetics 25), 129–163. Berlin & Boston, MA: De Gruyter Mouton.

Fischer, Kerstin. 2016. Robots as confederates: How robots can and should support research in the humanities. In Johanna Seibt, Marco Nørskov & Soren Schack Andersen (eds.), *What social robots can and should do*, 60–66. Amsterdam: IOS Press.

Grice, Martine & Stefan Baumann. 2002. Deutsche Intonation und GToBI. *Linguistische Berichte* 191, 267–298.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Harnad, Stevan. 1990. The symbol grounding problem. *Physica D: Nonlinear Phenomena* 42, 335–346.

Hawkins, Sarah. 2003. Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics* 31, 373–405.

Hedberg, Nancy, Juan Manuel Sosa & Lorna Fadden. 2004. Meanings and configurations of questions in English. *Proceedings 2nd International Conference of Speech Prosody*, Nanjing, Japan, 375–378.

Jones, Daniel. 1909. *Intonation curves: A collection of phonetic texts, in which intonation is marked throughout by means of curved lines on a musical stave*. Leipzig & Berlin: B. G. Teubner.

Kelly, Thomas F. 2014. *Capturing music: The story of notation*. New York: WW Norton & Company.

Kingdon, Roger. 1958. *The groundwork of English intonation*. London: Longmans.

Kohler, Klaus J. 2005. Timing and communicative functions of pitch contours. *Phonetica* 62, 88–105.

Kügler, Frank, Bernadett Smolibocki, Denis Arnold, Stefan Baumann, Bettina Braun, Martine Grice, Stefanie Jannedy, Jan Michalsky, Oliver Niebuhr, Jörg Peters, Simon Ritter, Christine T. Röhr, Antje Schweitzer, Katrin Schweitzer & Petra Wagner. 2015. DIMA: Annotation guidelines for German intonation. *18th International Congress of Phonetic Sciences* (ICPhS XVIII), Glasgow, Scotland, 317.

Ladd, D. Robert. 2008. *Intonational phonology*, 2nd edn. Cambridge: Cambridge University Press.

Lakoff, George. 1990. *Women, fire, and dangerous things*. Chicago, IL: University of Chicago Press.

Levinson, Stephen C. 2006. On the human 'Interaction Engine'. In N. J. Enfield & Stephen C. Levinson (eds.), *Roots of human sociality*, 39–69. New York: Berg.

Liberman, Alvin M. & Doug H. Whalen. 2000. On the relation of speech to language. *Trends in Cognitive Sciences* 4, 187–196.

Litman, Diane J. & Kate Forbes-Riley. 2006. Recognizing student emotions and attitudes on the basis of utterances in spoken tutoring dialogues with both human and computer tutors. *Speech Communication* 48, 559–590.

Mayer, Jörg. 1995. Transcription of German intonation: The Stuttgart System. Ms., University of Stuttgart.

Mayo, Catherine, Robert A. J. Clark & Simon King. 2011. Listeners' weighting of acoustic cues to synthetic speech naturalness: A multidimensional scaling analysis. *Speech Communication* 53, 311–326.

Melis, Alicia P., Patricia Grocke, Josefine Kalbitz & Michael Tomasello. 2016. One for you, one for me: Humans' unique turn-taking skills. *Psychological Science* 27, 987–996.

Niebuhr, Oliver. 2013a. The acoustic complexity of intonation. In Eva Liina Asu & Pärtel Lippus (eds.), *Nordic Prosody XI*, 15–29. Frankfurt: Peter Lang.

Niebuhr, Oliver. 2013b. Resistance is futile: The intonation between continuation rise and calling contour in German. *Proceedings of 14th International Interspeech Conference*, Lyon, France, 225–229.

Niebuhr, Oliver. 2015. Stepped intonation contours. In Radek Skarnitzl & Oliver Niebuhr (eds.), *Tackling the complexity in speech*, 39–74. Prague: Charles University Press.

Niebuhr, Oliver, Silke Tegtmeier & Alexander Brem. 2017. Advancing research and practice in entrepreneurship through speech analysis: From descriptive rhetorical terms to phonetically informed acoustic charisma metrics. *Journal of Speech Sciences* 6, 3–26.

Ogden, Richard. 2010. Prosodic constructions in making complaints. In Dagmar Barth-Weingarten, Elisabeth Reber & Margret Selting (eds.), *Prosody in interaction*, 81–104. Amsterdam: John Benjamins.

Ogden, Richard A. 2012. Prosodies in conversation. In Oliver Niebuhr (ed.), *Understanding prosody: The role of context, function, and communication*, 201–217. Berlin & New York: de Gruyter.

Peters, Jörg. 2014. *Intonation*. Winter: Heidelberg.

316 Prieto, Pilar. 2015. Intonational meaning. *Wiley Interdisciplinary Reviews: Cognitive Science* 6, 371–381.
317 Rao, Rajiv. 2013. Intonational variation in third party complaints in Spanish. *Journal of Speech Sciences*
318 3, 141–168.
319 Rathcke, Tamara, [V.] & Jonathan Harrington. 2010. The variability of early accent peaks in Standard
320 German. In Cécile Fougeron, Barbara Kühnert, Mariopaola D'Imperio & M. Nathalie Vallée (eds.),
321 *Laboratory Phonology*, vol. 10, 533–555. Berlin & New York: de Gruyter Mouton.
322 Rathcke, Tamara V. 2016. How truncating are 'truncating languages'? Evidence from Russian and
323 German. *Phonetica* 73, 194–228.
324 Rodero, Emma. 2017. Effectiveness, attention, and recall of human and artificial voices in an advertising
325 story: Prosody influence and functions of voices. *Computers in Human Behavior* 77, 336–346.
326 Sandry, Elanor. 2015. *Robots and communication*. Basingstoke: Palgrave Macmillan.
327 Selting, Margret, Peter Auer, Dagmar Barth-Weingarten, Jörg R. Bergmann, Pia Bergmann, Karin Birkner,
328 Elizabeth Couper-Kuhlen, Arnulf Deppermann, Peter Gilles, Susanne Günthner, Martin Hartung,
329 Friederike Kern, Christian Mertzlufft, Christian Meyer, Miriam Morek, Frank Oberzaucher, Jörg Peters,
330 Uta Quasthoff, Wilfried Schütte, Anja Stukenbrock & Susanne Uhmann. 2009. Gesprächsanalytisches
331 Transkriptionssystem 2 (GAT 2). *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* 10,
332 353–402.
333 Steele, Joshua. 1779. *Prosodia rationalis: Or, an essay towards establishing the melody and measure of*
334 *speech, to be expressed and perpetuated by peculiar symbols*. London: J. Nichols.
335 Ward, Nigel G. & David DeVault. 2016. Challenges in building highly-interactive dialog systems. *AI*
336 *Magazine* 37, 7–18.
337 Ward, Nigel G. & Paola Gallardo. 2017. Non-native differences in prosodic-construction use. *Dialogue &*
338 *Discourse* 8, 1–30.
339 Wichmann, Anne, Jill House & Toni Rietveld. 2000. Discourse constraints on f0 peak timing in English. In
340 Antonis Botinis (ed.), *Intonation: Analysis, modeling and technology*, 163–182. Dordrecht: Springer.
341 Wolters, Maria K., Christine Johnson, Pauline E. Campbell, Christine G. DePlacido & Brian McKinstry.
342 2014. Can older people remember medication reminders presented using synthetic speech? *Journal of*
343 *the American Medical Informatics Association* 22, 35–42.
344 Zen, Heiga, Andrew Senior & Mike Schuster. 2013. Statistical parametric speech synthesis using deep
345 neural networks. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal*
346 *Processing* (ICASSP), Vancouver, Canada, 7962–7966.