

Identifying Speakers' Personal Information in Phone Conversations

Shi Hu, Peter Lipay

Stanford University

{s3hu and plipay}@cs.stanford.edu

Abstract

This is the paper layout specification and template definition for the InterSpeech 2012 Conference, which will be held in Portland, Oregon, USA. This template has been generated from previous InterSpeech templates and is sized in A4 (ISO 216, 210 mm \times 297 mm). The format is essentially the one used for the IEEE ICASSP conferences. You must include index terms as shown below.

Index Terms: speech synthesis, unit selection, join costs

1. Introduction

This template can be found on the conference website. Please use either an OpenOffice, MS-Word® or a L^AT_EX format file when preparing your submission. Information for full paper submission is available on the conference web site.

2. Page layout and style

All papers must be submitted in compliance with the template as outlined. Please select the appropriate template from the author kit to develop your paper. Check details of your final PDF submission against the corresponding example file.

2.1. Basic layout features

- Proceedings will be created in A4 paper size. Authors must therefore submit their papers in A4 paper size.
- Two columns are used except for the title part and possibly for large figures that need a full page width.
- Left margin is 20 mm.
- Column width is 80 mm.
- Spacing between columns is 10 mm.
- Top margin 25 mm (except for the first page which is 30 mm to the title top).
- Text height (without headers and footers) is maximum 235 mm.

- Headers and footers must be left empty. The conference publication company will add information for publication.
- Check indentations and spacings by comparing to this example file (in PDF).

2.1.1. Headings

Section headings are centered in boldface with the first word capitalized and the rest of the heading in lower case. Sub-headings appear like major headings, except they start at the left margin in the column. Sub-sub-headings appear like sub-headings, except they are in italics and not boldface. See the examples given in this file. No more than 3 levels of headings should be used.

2.2. Text font

Times or Times Roman font is used for the main text. Font size in the main text must be 9 points, and in the References section 8 points. Other font types may be used if needed for special purposes. All fonts must be embedded during the formation of the final PDF.

L^AT_EX users: users should use Adobe Type 1 fonts such as Times or Times Roman. These are used automatically by the interspeech2012.sty style file. Authors must not use Type 3 (bitmap) fonts.

2.3. Figures

All figures must be centered on the column (or page, if the figure spans both columns). Figure captions should follow each figure and have the format given in Fig. 1.

Figures should preferably be line drawings. If they contain gray levels or colors, they should be checked to print well on a high-quality non-color laser printer.

Graphics (i.e. illustrations, figures) should not use stipple fill patterns because they may not reproduce properly. Please use only *solid* fill colors.

Figures which span two columns (i.e. occupy full page width) should be placed at the top or bottom of a page.

2.4. Tables

An example of a table is shown as Table 1. Somewhat different styles are allowed according to the type and purpose of the table. The caption text may be above or below the table.

Table 1: *This is an example of a table.*

ratio	decibels
1/1	0
2/1	≈ 6
3.16	10
10/1	20
1/10	-20
100/1	40
1000/1	60

2.5. Equations

Equations should be placed on separate lines and numbered. Examples of equations are given below. Particularly,

$$x(t) = s(f_\omega(t)) \quad (1)$$

where $f_\omega(t)$ is a special warping function

$$f_\omega(t) = \frac{1}{2\pi j} \oint_C \frac{\nu^{-1k} d\nu}{(1 - \beta\nu^{-1})(\nu^{-1} - \beta)} \quad (2)$$

A residue theorem states that

$$\oint_C F(z) dz = 2\pi j \sum_k \text{Res}[F(z), p_k] \quad (3)$$

Applying (3) to (1), it is straightforward to see that

$$1 + 1 = \pi \quad (4)$$

Finally we have proven the secret theorem of all speech sciences. No more math is needed to show how useful the result is!

2.6. Hyperlinks

For technical reasons, the proceedings editor will strip all active links from the papers during processing. Hyperlinks can be included in your paper, if written in full, e. g. "http://www.foo.com/index.html". The link text must be all black. Please make sure that they present no problems in printing to paper.

2.7. Multimedia files

The InterSpeech 2012 organizing committee offers the possibility to submit multimedia files. These files are meant for audio-visual illustrations that cannot be conveyed in text, tables and graphs. Just like you would

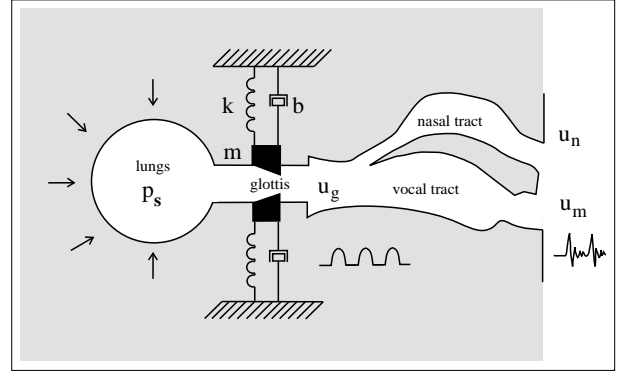


Figure 1: *Schematic diagram of speech production.*

when including graphics, make sure that you have sufficient author rights to the multimedia materials that you submit for publication. The proceeding media will *not* contain readers or players, so be sure to use widely accepted file formats and codecs, such as .mpg for video, .wav or .mp3 for audio, and .png or .jpg for images. The files you submit will be accessible from the abstract cards on the media and via a bookmark in the manuscript. From within the manuscript, refer to a multimedia illustration by its filename. Use short file names without blanks.

2.8. Page numbering

Final page numbers will be added later to the document electronically. *Don't make any footers or headers!*

2.9. References

The reference format is the standard IEEE one. References should be numbered in order of appearance, for example here [1], there [2], and everywhere [3]. Don't refer to citations as if they were parts of speech.

2.10. Abstract

The total length of the abstract is limited to 1000 characters. The abstract included in your paper and the one you enter during web-based submission must be identical. Avoid non-ASCII characters or symbols as they may not display correctly in the abstract book.

2.11. Author affiliation

Please list country names as part of the affiliation for each country.

2.12. Submitted files

Authors are requested to submit PDF files of their manuscripts. You can use commercially available tools or for instance <http://www.pdfforge.org/products/pdfcreator>, or pdflatex. The PDF file should comply with the following require-

ments: (a) there must be no password protection on the PDF file at all; (b) all fonts must be embedded; and (c) the file must be text searchable (do CTRL-F and try to find a common word such as 'the'). The proceedings editors will contact authors of non-complying files to obtain a replacement. In order not to endanger the preparation of the proceedings, papers for which a replacement is not provided timely will be withdrawn.

3. Conclusions

Authors must proof read their PDF file prior to submission to ensure it is correct. Authors should not rely on proofreading their source file. Please proofread the final PDF file before it is submitted.

4. Acknowledgements

The InterSpeech 2012 organizing committee would like to thank the organizing committee of InterSpeech 2008, 2009, 2010, 2011 for their help and for kindly providing the template files.

5. References

- [1] Smith, J. O. and Abel, J. S., "Bark and ERB Bilinear Transforms", IEEE Trans. Speech and Audio Proc., 7(6):697–708, 1999.
- [2] Soquet, A., Saerens, M. and Jospa, P., "Acoustic-articulatory inversion", in T. Kohonen [Ed], Artificial Neural Networks, 371-376, Elsevier, 1991.
- [3] Stone, H.S., "On the uniqueness of the convolution theorem for the Fourier transform", NEC Labs. Amer. Princeton, NJ. Online: <http://citeseer.ist.psu.edu/176038.html>, accessed on 19 Mar 2008.