

## SUBSCRIPTIONS

*Natural Language Engineering* (ISSN 1351-3249; electronic 1469-8110) is published six times a year in January, March, May, July, September and November. Six parts form a volume. The subscription price (excluding VAT and sales tax) of volume 25 (2019) (which includes postage) is £407 net (US\$664 in the USA, Canada and Mexico) for institutions print and electronic, institutions electronic only £337/\$555; £49 net (US\$79 in the USA, Canada and Mexico) for individuals ordering direct from the publisher and certifying that the journal is for their personal use. Single parts are £75 (US\$125 in the USA, Canada and Mexico) plus postage.

Orders, which must be accompanied by payment, may be sent to a bookseller, subscription agent or direct to the publisher: Cambridge University Press, Journals Fulfillment Department, UPH, Shaftesbury Road, Cambridge CB2 8BS, UK; or in the USA, Canada and Mexico: Cambridge University Press, Journals Fulfillment Department, 1 Liberty Plaza, Floor 20, New York, NY 10006, USA. EU subscribers who are not registered for VAT should add VAT at their country's rate. VAT registered subscribers should provide their VAT registration number. Japanese prices for institutions are available from Kinokuniya Company Ltd, PO Box 55, Chitose, Tokyo 156, Japan.

## COPYING

This journal is registered with the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, USA. Organizations in the USA who are also registered with C.C.C. may therefore copy material (beyond the limits permitted by sections 107 and 108 of U.S. Copyright law) subject to payment to C.C.C. of the per-copy fee of \$16.00. This consent does not extend to multiple copying for promotional or commercial purposes. Code 1351-3249/2015.

*ISI Tear Sheet Service*, 3501 Market Street, Philadelphia, PA 19104, USA, is authorised to supply single copies of separate articles for private use only.

Organisations authorised by the Copyright Licensing Agency may also copy material subject to the usual conditions.

*For all other use*, permission should be sought from Cambridge or from the American Branch of Cambridge University Press.

## INSTRUCTIONS FOR CONTRIBUTORS

Instructions for Contributors will be found on the Journal's webpage at [www.cambridge.org/nle](http://www.cambridge.org/nle). This journal and other journals are included in the Cambridge Journals Online service which can be found at [journals.cambridge.org](http://journals.cambridge.org)

© Cambridge University Press 2019

This journal issue has been printed on FSC-certified paper and cover board. FSC is an independent, non-governmental, non-for-profit organization established to promote the responsible management of the world's forests. Please see [www.fsc.org](http://www.fsc.org) for information.

# NATURAL LANGUAGE ENGINEERING

Volume 25 Part 1 January 2019

## CONTENTS

Editorial note <i>John Tait</i>	1
ARTICLES	
Arabic community question answering <i>Preslav Nakov, Lluís Màrquez, Alessandro Moschitti and Hamdy Mubarak</i>	5
A neural approach for inducing multilingual resources and natural language processing tools for low-resource languages <i>O. Zennaki, N. Semmar and L. Besacier</i>	43
Identifying signs of syntactic complexity for rule-based sentence simplification <i>Richard Evans and Constantin Orăsan</i>	69
Extractive multi-document summarization based on textual entailment and sentence compression via knapsack problem <i>Ali Naserasadi, Hamid Khosravi and Faramarz Sadeghi</i>	121
The effect of morphology in named entity recognition with sequence tagging <i>Onur Güngör, Tunga Güngör and Suzan Üsküdarlı</i>	147
Efficient Phrase Table pruning for Hindi to English machine translation through syntactic and marker-based filtering and hybrid similarity measurement <i>Niladri Chatterjee and Susmita Gupta</i>	171
Industry Watch	
Law and Word Order: NLP in Legal Tech <i>Robert Dale</i>	211

The cover illustration is computer generated from an engraving of the Tower of Babel in an eighteenth-century bible.

### Cambridge Core

For further information about this journal please go to the journal website at:  
[cambridge.org/nle](http://cambridge.org/nle)



MIX  
Paper from  
responsible sources  
FSC® C007785

**CAMBRIDGE**  
UNIVERSITY PRESS