

ICE
INTERNATIONAL CORPUS OF ENGLISH

**Manual for the
Trinidad and Tobago component
(ICE-T&T)**

MICHAEL WESTPHAL, DAGMAR DEUBER, KA MAN LAU

**ENGLISH DEPARTMENT
UNIVERSITY OF MÜNSTER, GERMANY**

2023

Table of contents

Part A: Background	1
1. ICE-T&T project team	1
2. Acknowledgements	2
3. Terms and conditions for the use of ICE-T&T	3
3.1 Definitions	3
3.2 General Terms and Conditions	3
4. Copyright information	4
5. English in Trinidad and Tobago	4
6. Publications related to ICE-T&T	6
Part B: ICE-T&T	7
7. Structure of ICE-T&T	7
8. Markup conventions	9
9. Spoken component	11
9.1 Editorial practice	11
9.2 Background information	15
10. Written component: Background information	16
Part C: References	17

Part A: Background

1. ICE-T&T project team

ICE-Trinidad & Tobago (ICE-T&T) was initiated in 2006 by Valerie Youssef, then Professor of Linguistics in the Department of Modern Languages and Linguistics, University of the West Indies, St. Augustine, Trinidad, and Dagmar Deuber, then a visiting postdoctoral researcher and lecturer in the same department; the project continued when the latter moved back to her then home university of Freiburg in 2007. After taking up the Chair of Variation Linguistics in the Department of English at the University of Münster in 2010, Dagmar Deuber coordinated the project until its completion in 2022.

Over the years, a large number of people from the three different institutions involved contributed to the compilation of ICE-T&T:

at the University of the West Indies in St. Augustine, Trinidad

- Ryan Durgasingh (also later at the University of Münster)
- Janielle Garcia
- Patrice Camille Quammie

at the University of Freiburg, Germany

- Glenda Leung
- Luminita Trasca
- Christine Wender

at the University of Münster, Germany

- Katrin Dieckmann
- Judith Evers
- Friederike Hahn
- Melanie Hankins
- Eva Canan Hänsel
- Johanna Hartmann
- Jutta Heuger
- Lisa Jansen
- Meike Jensen
- Darlene Ka Man Lau
- Bianca Melchien
- Dorit Neumann
- Jane Philip
- Carla Postmeier
- Carolina Rodríguez Beltrán
- Inga Schafmeister
- Lisha Tang
- Michael Westphal
- Guyanne Wilson

2. Acknowledgements

Our first acknowledgement goes to all the speakers and writers who have contributed to the project by participating in recordings or contributing written material to the corpus. A linguistic corpus is made up of a multitude of individual voices and contributions, which are often easily forgotten in quantitative analyses. However, this corpus was only made possible through the contributions of many individuals. We are deeply indebted to all of the individual contributors.

Many of the early texts for the corpus were recorded and transcribed by students at the University of the West Indies, St. Augustine, who participated in the class LING 3099 “Special Project in Linguistics” in the academic year 2006/07. We are also indebted to these contributions that helped to kick-start the project in the initial stages.

ICE-T&T received a further major impetus during its initial stages from two grants by the University of the West Indies’ Research and Publication Fund.

The compilation of the corpus would not have been possible without the support of many institutions in Trinidad and Tobago which made texts available, granted copyright permissions, or helped in many other ways. These institutions are listed individually in the accompanying background information files, which provide the background details and copyright information for each individual text.

Finally, we would like to specially thank Ryan Durgasingh, Glenda Leung, and Guyanne Wilson, who, as dedicated native speaker linguists, not only recorded, transcribed, and proofread a great number of texts but also contributed to the project in many more ways.

3. Terms and conditions for the use of ICE-T&T

By downloading the International Corpus of English – Trinidad and Tobago, the Licensee agrees as follows:

3.1 Definitions

- “Corpus” refers to the Trinidad and Tobago component of the International Corpus of English.
- “Institution” refers to the non-profit academic institution that the Licensee is a member of.
- “Licence” means the licence granted to the Licensee to use the Corpus in accordance with the terms and conditions of this agreement.

3.2 General Terms and Conditions

The Corpus must be used for non-profit academic research purposes only. The Licence cannot be transferred, lent, or re-sold.

The Licensee agrees not to reproduce or redistribute the Corpus or to use all or any part of the Corpus texts in any commercial product or service. A copy of the Corpus may be made for backup purposes.

Copyright in all texts in the Corpus is retained by the original copyright holders.

The Corpus may be fully installed onto the Licensee’s computer. The Licensee is allowed to make copies of the Corpus on computers within the Institution.

The Licence entitles all staff and students of the Institution to make use of the Corpus installed on the computers within the Institution.

It is the responsibility of the Licensee to ensure that the Corpus cannot be accessed from outside the Institution. The Licence does not entitle the Licensee to include the Corpus in a public-access internet site.

It is the responsibility of the Licensee to ensure that other users of the Corpus within the Institution are made aware of the terms of the Licence.

Publications based on the Corpus may include citations from texts only in a way which would be permitted under the fair dealings provision of copyright law.

All publications based on the Corpus must give credit to the International Corpus of English.

The Licensee agrees not to re-distribute the Corpus to any third party.

The general terms and conditions apply.

4. Copyright information

Copyright for the individual texts remains with the copyright holders as listed in the background information files along with a permission statement. Every effort has been made to obtain copyright permission from each individual copyright holder. In the case of a few publicly available texts we were not able to get in contact with the copyright holders, but we will be happy to amend the entries in question in a revised version if the relevant information is sent to us. Speakers who were recorded in person signed permission for the texts to be published in the corpus, while contributions of non-public written texts were made on condition that permission was granted for their inclusion in the corpus; in the background information files, this is indicated as “reproduced with permission from the speaker/writer“.

5. English in Trinidad and Tobago

Trinidad and Tobago form a twin-island state and are the two most south-eastern islands in the Caribbean archipelago. The islands were united as a British colony in 1889 and finally became independent from Britain in 1962. With approximately 1.4 million inhabitants (CIA, 2023) Trinidad and Tobago is the second most populous state in the anglophone Caribbean. While the two islands share English as the sole official language, the sociolinguistic situation in each has been shaped by diverging colonial histories.

Trinidad is the much bigger island (4,768 km²) and was first colonized by Spain starting in the late 15th century, which led to the devastation of the indigenous Amerindian population as well as their culture and language. From 1783 onwards, French-speaking settlers and enslaved people from other French-speaking islands came to Trinidad in large numbers. Hence, Trinidad turned into a French-speaking island, as French and Trinidadian French Creole dominated. By now, Trinidadian French Creole is nearly extinct. British colonization, starting in 1797, led to the anglicization of the island. The recruitment of new workers in the post-emancipation period (after 1833) – mainly as indentured laborers – from China, Madeira, Syria and above all from India made Trinidad an ethnically diverse island. The major ethnic division today is between Afro- and Indo-Trinidadians. However, many people self-identify as being of mixed heritage and there are small minorities of other ethnic groups. The anglicization of Trinidad led to the development of Trinidadian English Creole, an English-lexified Creole, and Standard Trinidadian English, the local standard variety.

In Trinidad today, Trinidadian English Creole is the first code for the majority of the population and most people acquire Standard Trinidadian English as a second dialect. Traditionally both codes serve distinct functions in society: Trinidadian English Creole dominates in informal private domains; it has a long history of stigmatization but also serves as a marker of solidarity and Trinidadian identity. In contrast, Standard Trinidadian English is targeted in formal public domains and is a marker of education and social status. However, Trinidadian English Creole has made inroads into domains formerly reserved for Standard English, and attitudes toward it have become more positive. Hence, the clear functional separation has become blurred. Linguistically, variation between Standard Trinidadian English and

Trinidadian English Creole is commonly modelled as a continuum. Trinidadian English Creole is classified as a mesolectal Creole, as the distance to Standard English is less pronounced than in other Caribbean islands, like Jamaica.

Several European powers tried to colonize Tobago, but their attempts failed due to the resistance of the indigenous Amerindian inhabitants and inter-European conflicts over the island. In 1763, Tobago became a British plantation colony and saw the arrival of larger numbers of British colonizers and enslaved people from West Africa. This phase of European colonization led to the fast decline and eventual extinction of the indigenous people of Tobago by the end of the 18th century. Tobago came under French control twice but was recaptured by the British and remained a British colony from 1803 until 1962. In contrast to Trinidad, no European power besides Britain had a lasting linguistic impact on Tobago, and the population of Tobago is primarily of African descent. The distinct colonial history of Tobago led to the emergence of a basilectal English Creole. This continues in use as a home variety (Youssef & James, 2004, p. 523) but the mesolectal English Creole shared by speakers on both islands has become more prominent due to inter-island movements, especially migration from Tobago to Trinidad, and the general spread of education (Youssef & James, 2004, p. 513).

For more detailed descriptions of the sociolinguistic context of English in Trinidad and Tobago see for example Youssef and James (2004, pp. 508–514), Deuber (2014, pp. 30–43), or Wilson (2014, pp. 7–14).

Language use sampled in ICE-T&T aims to represent Standard English usage in Trinidad and Tobago but also captures the influence of Creole on Standard English in the twin-island state. Following the ICE guidelines (Greenbaum, 1996), the corpus aims to represent speakers and writers over 18 who have received a formal education to the completion of secondary school. For a detailed definition of Standard English in the ICE project see Greenbaum (1996), and for a discussion of how the sociolinguistic situation of English in Trinidad and Tobago has been addressed in the compilation of ICE-T&T see Deuber (2010).

6. Publications related to ICE-T&T

- Deuber, D. (2009a). Caribbean ICE corpora: Some issues for fieldwork and analysis. In A. H. Jucker, M. Hundt, & D. Schreier (Eds.), *Corpora: Pragmatics and Discourse. Papers from the ICAME 29* (pp. 425–450). Rodopi.
- Deuber, D. (2009b). Standard English in the secondary school in Trinidad: Problems - properties - prospects. In T. Hoffmann & L. Siebers (Eds.), *World Englishes – Problems, properties and prospects: Selected papers from the 13th IAWC conference* (pp. 83–104). Benjamins. <https://doi.org/10.1075/veaw.g40.08deu>
- Deuber, D. (2010a). Modal verb usage at the interface of English and a related creole: A corpus-based study of can/could and will/would in Trinidadian English. *Journal of English Linguistics*, 38(2), 105–142. <https://doi.org/10.1177/0075424209348151>
- Deuber, D. (2010b). Standard English and situational variation: Sociolinguistic considerations in the compilation of ICE-Trinidad and Tobago. *ICAME Journal*, 34, 24–40.
- Deuber, D. (2014). *English in the Caribbean: Variation, style and standards in Jamaica and Trinidad*. Cambridge University Press.
- Deuber, D., Biewer, C., Hackert, S., & Hilbert, M. (2012). Will and would in selected New Englishes: General and variety-specific tendencies. In M. Hundt & U. Gut (Eds.), *Mapping unity and diversity in English world-wide: Corpus-based studies of New Englishes* (pp. 77–102). Benjamins. <https://doi.org/10.1075/veaw.g43.04deu>
- Deuber, D., Hänsel, E. C., & Westphal, M. (2021). Quotative be like in Trinidadian English. *World Englishes*, 40(3), 436–458. <https://doi.org/10.1111/weng.12465>
- Deuber, D., & Youssef, V. (2007). Teacher language in Trinidad: A pilot corpus study of direct and indirect creolisms in the verb phrase. In *Proceedings from the Corpus Linguistics 2007 Conference*.
- Neumaier, T. (2023). *Conversation in World Englishes: Turn-taking and cultural variation in Southeast Asian and Caribbean English*. Cambridge University Press.
- Westphal, M. (2021). Question tags across New Englishes. *World Englishes*, 40(4), 519–533. <https://doi.org/10.1111/weng.12538>
- Westphal, M., Lau, K. M., Hartmann, J., & Deuber, D. (2022). Phonetic variation in Standard English spoken by Trinidadian professionals. *Journal of Pidgin and Creole Languages*, 37(2), 357–394. <https://doi.org/10.1075/jpcl.00098.wes>
- Westphal, M., & Wilson, G. (2020). New Englishes, new methods: Focus on corpus linguistics. *Anglistik: International Journal of English Studies*, 31(3), 47–65. <https://doi.org/10.33675/ANGL/2020/3/7>
- Westphal, M., & Wilson, G. (2023). Creole and power: A critical discourse analysis of legal cross-examinations in ICE Jamaica and ICE Trinidad and Tobago. In G. Wilson & M. Westphal (Eds.), *New Englishes, new methods* (pp. 84–106). Benjamins. <https://doi.org/10.1075/veaw.g68.05wes>
- Wilson, G., Westphal, M., Hartmann, J., & Deuber, D. (2017). The use of question tags in different text types of Trinidadian English. *World Englishes*, 36(4), 726–743. <https://doi.org/10.1111/weng.12247>
- Youssef, V., & Deuber, D. (2007). ICE Trinidad and Tobago: Teacher language investigation in a university research class. In *Proceedings from the Corpus Linguistics 2007 Conference*.

Part B: ICE-T&T

7. Structure of ICE-T&T

ICE-T&T includes: the corpus itself (500 txt files), two background information spreadsheets (*ICE-T&T SPOKEN*; *ICE-T&T WRITTEN*) providing details of the texts as well as biographical data of the speakers and writers, and two further spreadsheets (*ICE-T&T SPOKEN word count*; *ICE-T&T WRITTEN word count*) that give the precise word count for each text.

The texts in ICE-T&T cover a time span of 10 years. The first recordings for the spoken component were carried out in 2006 and the final ones in 2013. ICE-T&T also covers public (written and spoken) materials with publications ranging from 2006 to 2015. Details on the time of recording or the publication of public (written and spoken files) are indicated in the two background information spreadsheets.

ICE-T&T closely follows the general ICE design described in Nelson (1996a). Overall, ICE-T&T includes 300 spoken and 200 written texts from different public and private text types. Table 1 shows an overview of the different text types included in ICE-T&T, the number of texts in each text type, and the systematic names of texts used for the naming of txt files. Each text contains approximately 2,000 words. The spreadsheets *ICE-T&T SPOKEN word count* and *ICE-T&T WRITTEN word count* provide the exact number of words for each text. Many texts consist of several subtexts of the same text type, which are grouped together into one text of about 2,000 words. Individual subtexts are marked with <I>...</I> in the files (see Section 8) and the background information spreadsheets provide information on the subtexts.

Table 1. Overview ICE text types

Spoken (300)	Dialogues (180)	Private (100)	Conversations (90) Phonecalls (10)	S1A-001 to 090 S1A-091 to 100
		Public (80)	Class Lessons (20) Broadcast Discussions (20) Broadcast Interviews (10) Parliamentary Debates (10) Legal cross-examinations (10) Business Transactions (10)	S1B-001 to 020 S1B-021 to 040 S1B-041 to 050 S1B-051 to 060 S1B-061 to 070 S1B-071 to 080
	Monologues (120)	Unscripted (70)	Commentaries (20) Unscripted Speeches (30) Demonstrations (10) Legal Presentations (10)	S2A-001 to 020 S2A-021 to 050 S2A-051 to 060 S2A-061 to 070
		Scripted (50)	Broadcast News (20) Broadcast Talks (20) Non-broadcast Talks (10)	S2B-001 to 020 S2B-021 to 040 S2B-041 to 050
Written (200)	Non-printed (50)	Student Writing (20)	Student Essays (10) Exam Scripts (10)	W1A-001 to 010 W1A-011 to 020
		Letters (30)	Social Letters (15) Business Letters (15)	W1B-001 to 015 W1B-016 to 030
	Printed (150)	Academic (40)	Humanities (10) Social Sciences (10) Natural Sciences (10) Technology (10)	W2A-001 to 010 W2A-011 to 020 W2A-021 to 030 W2A-031 to 040
		Popular (40)	Humanities (10) Social Sciences (10) Natural Sciences (10) Technology (10)	W2B-001 to 010 W2B-011 to 020 W2B-021 to 030 W2B-031 to 040
		Reportage (20)	Press reports (20)	W2C-001 to 020
		Instructional (20)	Administrative Writing (10) Skills/hobbies (10)	W2D-001 to 010 W2D-011 to 020
		Persuasive (10)	Editorials (10)	W2E-001 to 010
		Creative (20)	Novels (20)	W2F-001 to 020

8. Markup conventions

ICE-T&T uses the general markup conventions of the ICE project as detailed in Nelson (1996b, 2002a, 2000b). Tables 2 and 3 provide an overview of the markup conventions for the spoken and written component, respectively. For a detailed description of the ICE markup with examples see Nelson (1996b, 2002a, 2000b).

In contrast to ICE-Jamaica (Rosenfelder et al., 2009), where Jamaican Creole forms are marked as indigenous (i.e. <indig> ... </indig>), morpho-syntactic forms or lexical items typical of Trinidadian English Creole are not marked as indigenous. The indigenous markup is used for words from languages other than Trinidadian English/Creole that are or once were spoken on the islands – Amerindian languages, Spanish, French/Trinidadian French Creole, African languages such as Yoruba, Hindi/Bhojpuri.

Table 2. Markup conventions for spoken text: Overview

Markup symbol	Meaning
<\$A>, <\$B>, etc	Speaker identification
<l>...</l>	Subtext marker
<#>	Text unit marker
<O>...</O>	Untranscribed text
<?>...<?>	Uncertain transcription
<->...</->	Normative deletion
<+>...</+>	Normative insertion
<=>...</=>	Original normalization
<. >...</.>	Incomplete word
<}>...</}>	Normative replacement
<[>...</[>	Overlapping string
<{>...</{>	Overlapping string set
<, >	Short pause
<, , >	Long pause
<(>...</(>	Discontinuous word
<) >...</) >	Normalized disc. word
<X>...</X>	Extra-corpus text
<&>...</&>	Editorial comment
<@>...</@>	Changed name or word
<w>...</w>	Orthographic word
<quote>...</quote>	Quotation
<mention>...</mention>	Mention
<foreign>...</foreign>	Foreign word(s)
<indig>...</indig>	Indigenous word(s)
<unclear>...</unclear>	Unclear word(s)

Table 3. Markup conventions for written texts: Overview

Markup symbol	Meaning
<#>	Text unit marker
<l>...</l>	Subtext marker
<l>	Linebreak marker
<p>...</p>	Paragraph marker
<h>...</h>	Heading
<w>...</w>	Orthographic word
<X>...</X>	Extra-corpus text
<?>...</?>	Uncertain transcription
<O>...</O>	Untranscribed text
<. >...</.>	Incomplete word
<->...</->	Normative deletion
<+>...</+>	Normative insertion
<=>...</=>	Original normalization
<}>...</}>	Normative replacement
<&>...</&>	Editorial comment
<(>...</(>	Discontinuous word
<)>...</)>	Normalized disc. word
<@>...</@>	Changed name or word
<sb>...</sb>	Subscript
<sp>...</sp>	Superscript
...	Underline
<it>...</it>	Italics
<bold>...</bold>	Boldface
<typeface>...</typeface>	Change of typeface
<roman>...</roman>	Roman type
<smallcaps>...</smallcaps>	Small capitals
<footnote>...</footnote>	Footnote
<fnr>...</fnr>	Reference to footnote
<space>	Orthographic space
<quote>...</quote>	Quotation
...	Deleted text
<marginalia>...</marginalia>	Marginalia
<mention>...</mention>	Mention
<indig>...</indig>	Indigenous word(s)
<foreign>...</foreign>	Foreign word(s)

9. Spoken component

9.1 Editorial practice

Generally, British English spelling conventions are used, with *-ize* rather than *-ise*, i.e.:

- *-ize* (e.g. *minimize, organize*)
- *-izer* (e.g. *colonizer*)
- *-ized* (e.g. *socialized, vocalized*)
- *-izing* (e.g. *traumatizing, capitalizing*)
- *-ization* (e.g. *internationalization, organization*)
- but: *analyse*

Spelling of titles:

- *Dr*
- *Miss* (form of address without a name)
- *Mr*
- *Mrs*
- *Ms*
- *Sir*

The following Creole spellings are used in case the word in question is pronounced as indicated by the spelling:

- *aks* 'ask' (Winer, 2009)
- *likkle* 'little' (Winer, 2009)

Creole negator (also question particle):

- *ain't* (alternative forms according to Winer (2009, pp. 11, 325): *ain, ehn, en, ent*)

Discourse markers and filled pauses:

- *all right*
- *eh* (alternative form according to Winer (2009, p. 325): *ey*)
- *hm, hmm, hmmm, hmph, mh, mhm*
- *nah* (alternative form according to Winer (2009, p. 623): *na*)
- *OK*
- *uh, uhm*

Generally, spelling conventions for specific words in the spoken part are based on the *Oxford English Dictionary* (n.d.) and Winer (2009). All numbers (e.g. *twenty-three*) and double names (e.g. *Persad-Bissessar*) are hyphenated. Tables 4, 5, and 6 provide details on the spelling conventions for hyphenated words, single words, or separate words.

Table 4. Hyphenated words

[number]-year-old (adj)	<i>easy-going</i>	<i>left-hand</i> (attributive adj)	<i>politically-minded</i> (attributive)
<i>about-face</i>	<i>eco-adventure/-tourism</i>	<i>level-headed</i>	<i>Port-au-Prince</i>
<i>acquisition-</i>	<i>E-coli</i>	<i>line-out/-up</i> (noun)	<i>post-</i> and <i>pre-</i> as

<i>learning hypothesis</i>			prefixes (e.g. <i>post-Carnival</i> , <i>post-emancipation</i> , <i>pre-emptive</i> , <i>pre-/post-operative</i> , <i>pre-packaged</i> , <i>pre-qualification</i> , <i>pre-work</i>) but <i>postpaid</i> , <i>preheat</i> , <i>prepaid</i> , <i>preschool</i>
<i>Afro-Trinidadian</i>	<i>English-based/-speaking</i>	<i>lip-reader</i>	
<i>agro-farmers/-farming/-processing/-tech</i>	<i>e-Teck</i>	<i>lip-service</i>	
<i>air-condition</i>	<i>e-tutor</i>	<i>long-standing</i> (adj)	
<i>all-inclusive</i>	<i>ex-lover/-member</i>	<i>long-term</i>	
<i>all-purpose</i>	<i>extra-curricular</i>	<i>look-up</i> (adj)	<i>post-mortem</i>
<i>all-rounder</i>	<i>far-de-lance</i>	<i>loud-hailer</i>	<i>pro-</i> as prefix (e.g. <i>pro-slavery</i> , <i>pro-Syrian</i>)
<i>all-time</i>	<i>fer-de-lance</i>	<i>low-hanging</i>	<i>push-up</i>
<i>al-Muslimeen</i>	<i>finger-spinner</i>	<i>low-wage</i>	<i>quasi-</i> as prefix (e.g. <i>quasi-political</i>)
<i>anti-</i> as prefix (e.g. <i>anti-colonial</i> , <i>anti-slavery</i>)	<i>first-/second-/third-generation</i> (attributive)	<i>machine-readable</i>	<i>quito-quito</i>
<i>a-plenty</i>	<i>first-class</i> (adj)	<i>maintenance-wise</i>	<i>rah-rah-rah</i>
<i>aqua-centre</i>	<i>first-hand</i> (adj)	<i>make-word</i> (noun)	<i>re-</i> as prefix followed by <e> (e.g. <i>re-enactment</i> , <i>re-enter</i> , <i>re-examination</i>)
<i>at-risk</i> (adj)	<i>five-seater</i>	<i>man-of-war</i>	<i>red-bellied</i>
<i>attorney-at-law</i>	<i>flip-flop</i>	<i>mark-up</i> (noun)	<i>red-card</i> (verb)
<i>audio-visual</i>	<i>follow-up</i> (noun)	<i>maxi-taxi</i>	<i>left-/right-hander</i>
<i>aye-aye</i>	<i>fore-parents</i>	<i>medium-sized</i>	<i>Saint-Domingue</i>
<i>bad-talking</i>	<i>French-speaking</i>	<i>medium-term</i>	<i>same-sex</i> (adj)
<i>barrister-at-law</i>	<i>full-blown/-time/-scale</i>	<i>merry-go-round</i>	<i>second-hand</i>
<i>-based</i> (in compounds) (e.g. <i>English-based</i> , <i>faith-based</i> , <i>software-based</i>)	<i>fund-raising</i>	<i>metropolitan-like</i>	<i>self-</i> as prefix (e.g. <i>self-defense</i> , <i>self-government</i> , <i>self-love</i> , <i>self-service</i>)
<i>bean-counter</i>	<i>get-go</i>	<i>micro-enterprise</i>	<i>sell-out</i> (noun)
<i>beh-beh</i>	<i>gift-giving</i>	<i>micro-financial</i>	<i>semi-</i> as prefix (e.g. <i>semi-final</i> , <i>semi-standard</i>)
<i>bling-bling</i>	<i>go-ahead</i> (noun)	<i>mid-fifties</i>	<i>sentence-initially</i>
<i>blow-dry</i>	<i>god-daughter</i>	<i>mid-July</i>	<i>set-up</i> (noun)
<i>blue-backed</i>	<i>god-given</i>	<i>mid-nineties</i>	<i>shoot-out</i> (noun)
<i>bow-wow</i>	<i>golden-headed</i>	<i>mid-off</i>	<i>short-lived</i>
<i>boy-hand</i>	<i>good-/old-/young-looking</i> , etc.	<i>mid-on</i>	<i>short-term</i>
<i>brain-dead</i>	<i>go-round</i> (noun)	<i>mid-term</i>	<i>side-on</i> (adj, adv)

<i>break-in</i> (noun)	<i>government-appointed</i>	<i>mis-reception</i>	<i>single-minded</i> (adj)
<i>break-up</i> (noun)	<i>governor-general</i>	<i>modern-day</i>	<i>sit-up</i> (noun)
<i>British-born/-ruled</i> (adj)	<i>great-grandmother</i>	<i>brother-/mother-/niece-/sister-in-law</i>	<i>sky-scraper</i>
<i>built-in</i>	<i>half-line</i>	<i>much-needed</i>	<i>smoke-free</i>
<i>buy-out</i> (noun)	<i>hands-on</i>	<i>multi-purpose</i>	<i>so-called</i>
<i>bye-bye</i>	<i>hand-to-mouth</i>	<i>multi-storey</i>	<i>socio-cultural, socio-economic</i>
<i>by-products</i>	<i>hangers-on</i>	<i>must-see</i>	<i>soft-spoken</i>
<i>Cap-de-Ville</i>	<i>hardware-based</i>	<i>neo-liberal</i>	<i>sous-chef</i>
<i>capital-intense</i>	<i>hard-hitting/-working</i>	<i>neo-realistic</i>	<i>sou-sou</i>
<i>carnival-like</i> (adj)	<i>have-not</i>	<i>nerve-wracking</i>	<i>spin-off</i> (adj, noun)
<i>carry-on</i> (noun)	<i>heavy-weight</i>	<i>never-ending</i>	<i>spur-of-the-moment</i> (attributive)
<i>-centred</i> (in compounds) (e.g. <i>father-centred, student-centred</i>)	<i>high-flying</i> (adj)	<i>next-door</i> (attributive)	<i>stand-by</i>
<i>chac-chac</i>	<i>high-pitch</i>	<i>Niger-Congo</i>	<i>start-up</i> (company)
<i>change-over</i> (noun)	<i>high-ranking</i>	<i>nine-wicket</i> (attributive)	<i>state-of-the-art</i>
<i>check-up</i> (noun)	<i>hip-hop</i>	<i>no-case</i> (attributive)	<i>sub-Saharan</i>
<i>chewing-gum</i>	<i>hoity-toity</i>	<i>non-</i> as prefix (adj, noun) (e.g. <i>non-interference, non-standard, non-white</i>)	<i>take-home</i> (noun)
<i>child-mother</i>	<i>home-mates</i>	<i>no-no</i> (adj, noun)	<i>talked-about</i>
<i>Christ-like</i>	<i>horse-racing</i>	<i>no-nonsense</i>	<i>tea-time</i>
<i>clean-up</i> (noun)	<i>hyper-correct</i>	<i>no-parking</i> (attributive)	<i>thank-you</i> (attributive)
<i>Coca-Cola</i>	<i>ill-conceived</i>	<i>north-west, etc.</i>	<i>The Inter-Religious Organization of Trinidad and Tobago</i>
<i>cold-blooded</i>	<i>in-depth</i>	<i>Nur-E-Islam</i>	<i>three-pointer</i>
<i>commander-in-chief</i>	<i>Indo-Trinidadian</i>	<i>off-</i> as prefix with noun (adj) (e.g. <i>off-campus, off-route</i>)	<i>three-quarter/s</i>
<i>common-law</i>	<i>in-fighting</i>	<i>energy-/oil-/tourism-dependent</i>	<i>throw-in</i> (noun) (soccer/football)
<i>computer-based</i>	<i>in-service</i> (adj)	<i>on-camera</i> (adj)	<i>top-down</i> (adj)
<i>co-operated/-owned</i>	<i>Inter-American</i>	<i>one-on-one</i>	<i>T-shirt</i>
<i>Creole-based/-like</i> (adj)	<i>inter-hall</i>	<i>one-way</i> (adj, noun)	<i>turn-out</i> (noun)
cricket terms:	<i>island-wide</i>	<i>on-side</i> (noun)	<i>turn-taking</i>

<i>mid-off, off-drive, off-spinner, play-off, run-off (but: offhand, offside, off stump)</i>			
<i>crop-over</i>	<i>jam-packed</i>	<i>open-heart</i> (attributive)	<i>two-acre/-lane/-story/-way/-week</i> (attributive)
<i>cross-comparison</i>	<i>kick-off</i> (noun)	<i>open-minded</i>	<i>user-friendly</i>
<i>cross-country</i>	<i>kilowatt-hour</i>	<i>open-source</i>	<i>US-led</i>
<i>cross-cultural</i>	<i>knife-like</i>	over- as prefix with verbs (+past participles) and nouns (e.g. <i>over-consumption, over-qualified, over-reliance</i>)	vice- as prefix (e.g. <i>vice-president, vice-principal</i>)
<i>cross-examine/-examination</i>	<i>know-how</i>	<i>pan-American/-Caribbean/-Trinbago</i>	<i>vis-à-vis</i>
<i>cut-off</i> (noun)	<i>knowledge-based</i>	-fried as suffix (e.g. <i>deep-fried, pan-fried</i>)	<i>waist-belt</i>
<i>decision-making</i>	<i>Kool-Aid</i>	<i>part-time</i>	well- as prefix (e.g. <i>well-being, well-deserved, well-known</i>)
<i>dhal-belly</i>	<i>laid-back</i>	<i>Patrick Manning-led</i>	-wise as suffix (e.g. <i>business-wise</i>)
<i>disease-prone/-resistant</i>	<i>large-scale</i>	<i>play-off</i> (noun)	<i>world-wide</i>
<i>doo-doo</i>	<i>law-making</i>	<i>Point-à-Pierre, Pointe-à-Pierre</i>	<i>worst-case</i> (adj)
<i>double-sided</i>	<i>lead-up</i> (noun)	<i>point-blank</i>	<i>x-ray</i>

Table 5. Single words

<i>babysit</i>	<i>legside</i>	<i>nationwide</i>	<i>subcommittee</i>
<i>bicarbonate</i>	<i>maladministration</i>	<i>neoclassical</i>	<i>subculture</i>
<i>bilingual</i>	<i>malintent</i>	<i>offhand, offside</i>	<i>subpopulation</i>
<i>canerow plaits</i>	<i>micropolitics</i>	<i>over/underextension</i>	<i>subsystemic</i>
<i>counterplot</i>	<i>midwicket</i>	<i>ownish</i>	<i>sunscreen</i>
<i>decentralize*</i>	<i>mindset</i>	<i>policymaker</i>	<i>threefold</i>
<i>email</i>	<i>multidimensional</i>	<i>postpaid, preheat, prepaid, preschool</i>	<i>unnormal</i>
<i>frontliner</i>	<i>multimillionaire</i>	re- as prefix (e.g. <i>recheck,</i>	<i>upfield</i>
<i>handover</i> (noun)	<i>multivariate</i>	<i>rechristened,</i>	
<i>landowning</i>	<i>namesake</i>	<i>reorder</i>)	

Table 6. Separate words

<i>bling* out</i>	<i>customer oriented</i>	<i>high paced</i>	<i>off stump</i>
<i>brass tacks</i>	<i>cutting edge</i> (adj)	<i>high school</i>	<i>open source</i>
<i>British style</i>	<i>et cetera</i>	<i>ice cream</i>	<i>per cent</i>
<i>busy busy</i> (reduplication)	<i>evidence in chief</i>	<i>middle stump</i>	<i>role model</i>
<i>cock a doodle doo</i>	<i>far away/ side/ off</i>	<i>multi arts</i>	<i>sugar cane</i>
<i>coffee growing</i>	<i>fast paced/ tracks</i>	<i>multi sector</i>	
<i>cool down</i> (noun)	<i>four lane</i>	<i>north coast</i>	
<i>curl out</i> (noun)	<i>heads up</i>	<i>old time</i> (adj)	

9.2 Background information

The spreadsheet *ICE-T&T SPOKEN* gives the following information for each text (as available/applicable): text type category, text code of the specific text, subtexts, number of participants, speaker letter codes (i.e. A, B, C, ..., Z), date, time, and place of recording, title/topic, channel, organising body, audience and audience size as well as copyright permission. For the speakers the following biographical information is included (if available): nationality, first language, age, gender, ethnicity, Trinidadian/Tobagonian, educational level, education, occupation, and other speaker details. Speaker names are included if known unless a speaker wished to be anonymized. In the latter case, their first name and/or surname has been replaced with an anonymized name if it appeared in the text and the anonymized name is also given in the spreadsheet (preceded by @). Otherwise speakers who wished to be anonymized are listed as ANON(NUMBER) in the spreadsheet (if a speaker appears in more than one text they retain the same number).

Overall, the spoken component of ICE-T&T is made up of contributions from 1128 speakers.

The corpus contains contributions from Trinidadian and Tobagonian speakers. Information on whether speakers self-identified as Trinidadian or Tobagonian is provided consistently in the spreadsheet only for direct conversations (S1A-001 to S1A-090). Texts S1A-001 to S1A-058 are conversations among Trinidadians. S1A-059 to S1A-073 are mixed conversations with Trinidadian as well as Tobagonian speakers. S1A-074 to S1A-090 are Tobagonians-only conversations. Information on whether speakers self-identified as Trinidadian, Tobagonian, or Trinidadian/Tobagonian is available for 381 speakers. Overall, 267 (70.1%) speakers self-identified as Trinidadian, 110 (28.9%) as Tobagonian, and 4 (1.0%) as Trinidadian/Tobagonian.

Information on gender is available for 1092 speakers: 494 (54.8%) speakers are female; 494 (45.2%) speakers are male.

For the operationalization of speakers' age, 4 age groups are used: 18-25, 26-45, 46-65, and 66+. Information on age is available for 637 speakers. Table 7 provides an overview of the age distribution of these speakers.

Table 7. Age groups (N=637)

age group	18-25	26-45	46-65	66+
speakers (%)	159 (25.0)	340 (53.4)	124 (19.5)	14 (2.2)

The major categories for speakers' (mostly self-reported) ethnicity are Afro-, Indo-, and mixed-Trinidadian/Tobagonian; other ethnicities include Indigenous-, Syrian-, and white-Trinidadian/Tobagonian. Information on ethnicity is available for 746 speakers. Table 8 provides an overview of these speakers' ethnicity.

Table 8. Ethnicity (N=746)

ethnicity	Afro	Indo	Mixed	other
speakers (%)	340 (45.6)	227 (30.4)	117 (15.7)	62 (8.3)

10. Written component: Background information

The spreadsheet *ICE-T&T WRITTEN* gives the following information for each text (as available/applicable): text type category, text code of the specific text, subtexts, and copyright permission; for printed texts: date of publication, place of publication, main title (title of article/chapter), editor(s), source title, source type, number of periodical, pages, publisher, publisher's (e-mail) address, ISBN, ISSN, URL, retrieval date, copyright holder, and scope; for non-printed texts: date of writing, format, title/topic/subject and organising body and course (for student writing). For the writers the following biographical information is included (if available): nationality, age, gender, ethnicity, educational level, education, occupation, affiliations, and other writer details. Writers' names are given for printed texts, anonymized names or ANON(NUMBER) for non-printed texts (cf. section 9.2 above).

Overall, the written component of ICE-T&T is made up of contributions from 488 writers.

Information on gender is available for 415 writers: 251 (60.5%) writers are female; 164 (39.5%) writers are male.

The same four age groups are used for the speakers (see Section 9.2) and writers. Information on age is available for 150 writers. Table 9 provides an overview of the age distribution of these speakers.

Table 9. Age groups (N=150)

age group	18-25	26-45	46-65	66+
writers (%)	31 (20.7)	83 (55.3)	31 (20.7)	5 (3.3)

The same categories for ethnicity are used for speakers (see Section 9.2) and writers. Information on ethnicity is available for 132 writers. Table 10 provides an overview of these speakers' ethnicity.

Table 10. Ethnicity (N=132)

ethnicity	Afro	Indo	Mixed	other
writers (%)	50 (37.9)	62 (47.0)	17 (12.9)	3 (2.3)

Part C: References

- CIA. (2023, March 28). *Trinidad and Tobago*. <https://www.cia.gov/the-world-factbook/countries/trinidad-and-tobago/>
- Deuber, D. (2010). Standard English and situational variation: Sociolinguistic considerations in the compilation of ICE-Trinidad and Tobago. *ICAME Journal*, 34, 24–40.
- Deuber, D. (2014). *English in the Caribbean: Variation, style and standards in Jamaica and Trinidad*. Cambridge University Press.
- Greenbaum, S. (1996). Introducing ICE. In S. Greenbaum (Ed.), *Comparing English worldwide: The International Corpus of English* (pp. 3–12). Clarendon Press.
- Nelson, G. (1996a). The design of the corpus. In S. Greenbaum (Ed.), *Comparing English worldwide: The International Corpus of English* (pp. 27–35). Clarendon Press.
- Nelson, G. (1996b). Markup systems. In S. Greenbaum (Ed.), *Comparing English worldwide: The International Corpus of English* (pp. 36–53). Clarendon Press.
- Nelson, G. (2002a). *International Corpus of English: Markup manual for spoken texts*. Retrieved from <https://www.ice-corpora.uzh.ch/en/manuals.html>
- Nelson, G. (2002b). *International Corpus of English: Markup manual for written texts*. Retrieved from <https://www.ice-corpora.uzh.ch/en/manuals.html>
- Oxford English Dictionary. (n.d.). *OED online*. Retrieved from <https://www.oed.com/>
- Rosenfelder, I., Jantos, S., Höhn, N., & Mair, C. (2009). *International Corpus of English: Manual for the Jamaican component (ICE-JA)*. Department of English, University of Freiburg.
- Wilson, G. (2014). *The sociolinguistics of singing: Dialect and style in classical choral singing in Trinidad* [Doctoral dissertation, Westfälische Wilhelms-Universität Münster]. miami: Publication server of the University of Münster. https://repositorium.uni-muenster.de/document/miami/bf42529f-fc52-48fb-9678-0b1aee851c4c/diss_wilson_buchblock.pdf
- Winer, L. (2009). *Dictionary of the English/Creole of Trinidad and Tobago*. McGill-Queen's University Press.
- Youssef, V., & Winford, J. (2004). The creoles of Trinidad and Tobago: Phonology. In E. W. Schneider, K. Burrige, B. Kortmann, R. Mesthrie, & C. Upton (Eds.), *A handbook of varieties of English: Volume 1: Phonology* (pp. 508–524). Mouton de Gruyter. <https://doi.org/10.1515/9783110197181-035>