Josai International University Bulletin Vol. 34, No. 2, March 2026　25-40

〈Research articles〉

# Twenty Ten or Two Thousand Ten?
## Expressing the Year in English

Laura Huston

Abstract

Unlike previous year numerals, the year numeral 2010 has more than one verbalization option in spoken English. The author eventually labels these options Pattern α (Alpha): *twenty ten*, Pattern β (Beta): *two thousand ten*, and Pattern γ (Gamma): *two thousand and ten*. The investigator finds a gap in corpus research: no peer-reviewed report can be found that counts the spoken instances of these verbalization patterns of the year numeral 2010 in English. The researcher uses audio recordings and transcripts of a long-running radio show and podcast, *This American Life*, as a corpus of convenience and finds 168 utterances, mostly by speakers of North American English, of the year numeral 2010. The most frequently used verbalization pattern proves to be Pattern β (102 utterances), followed by Pattern α (51) and Pattern γ (16). Analysis shows that male speaker preference for Pattern β is stronger than that same preference among female and non-binary speakers, at 65.83% versus 47.92%.

Key words: numerals, corpus study, podcast corpus

## Introduction

Written Arabic numbers are a stable, dependable source of meaning that does not change from one written language to the next, with the exception of such typographical features as decimal separators. For example, the amount “eight million six hundred seventy-five thousand three hundred nine and five hundredths” may be written in Japan as 8,675,309.05, in Indonesia as 8.675.309,05, and in Russia as 8 675 309,05. However, in the spoken domain, a number of three or more digits might be encoded or verbalized in a way that makes it more difficult for a new learner to understand.

Names for numbers are called *numerals* in linguistics. Besides numbers for counting generalized quantities (e.g. *twelve, a hundred, three thousand and eight, fifteen thousand*) and fractions of quantities (e.g. *two thirds*, *a [noun phrase] and a half*), numerals include expressions for such specialized quantities as money and telling the time. In the spoken domain, the meanings of some of these expressions

might not be immediately apparent to a learner of English. An English learner proficient in counting might hear "three twenty," "a dollar fifty," "a quarter to eleven," and "seven oh nine," but not realize that these spoken expressions mean £3.20, $1.50, 10:45 or 22:45, and either £7.09, $7.09, 07:09, or 19:09, depending on context. Numerals used for identification purposes can also be difficult for learners' listening comprehension. This is due to the way digits are "chunked," or grouped, when a multi-digit number is verbalized for speaking purposes in English. An English learner who easily understands the written expression "Room 213" might not expect to hear 213 verbalized as "two thirteen," or might not know how to group digits when reading a postal code aloud: does 46410 end in "one zero," or in "ten" ?

This heterogeneity for digit-chunking practices in the English verbalization of a numeral has emerged in the context of expressing the year numeral in English. This is of special importance for English learners whose L1, like Japanese, encodes year numerals no differently from any four-digit number, designating numeral strings as year numbers through some other means, such the Japanese use of *-nen* ("year") as a suffix after a year numeral. The L1 Japanese learner of English cannot listen for a suffix to identify a spoken string of numerals as a year number, but must learn to encode the string as soon as she hears it, and identify it as a year, and not, e.g., a clock reading or an amount of currency.

For four-digit year numerals in English, encoding is mostly regular: divide the four-digit number into two halves, and verbalize both these halves as two-digit numbers, e.g. the year 1066 is "ten sixty-six," and the year 1998 is "nineteen ninety-eight." There are exceptions to this rule. When the last two digits of a year number are both zeros, the zeros are read as "hundred," not as "zero zero." The year 1500 is "fifteen hundred," not "fifteen zero zero." When only the third digit of a year numeral is zero, that zero is read as "oh" in English. The year 1803 is verbalized in English as "eighteen oh three."

English verbalization of the year numerals 2000 to 2009 broke from the pattern described above. These year numerals are verbalized no different from counting numbers written with the same four digits. Then it became time to refer to the year 2010 in English speech and new options arose. An English-proficient listener accepts 3 different verbalizations of the year numeral 2010: *twenty ten, two thousand ten*, and *two thousand and ten*. Which verbalization pattern can a listener expect to encounter most often in the spoken domain of English? Are there any demographic patterns that align with these 3 verbalizations? This researcher was surprised to find no peer-reviewed articles on the topic and saw a gap and determined to fill it with data from a long-running radio show and podcast, *This American Life* (www.thisamericanlife.com).

## Literature Review

This paper investigates what words English-speakers use to indicate the year which is written *2010* in Arabic numbers. In any use of written Arabic numbers, no matter the language, there is no alternative to these four digits in this order to express the year 2010. However, in English, when reading this year numeral aloud (or expressing the idea of this year numeral, whether looking at the written digits or not), there is more than one set of words available to choose from. That is, there is more than one way to put into words, to *verbalize,* the year numeral 2010 in English.

Olsson (1988) described different ways to verbalize numerals in various languages, including English. He used the term *maximal fusing* to mean the use of the number words, or numerals, expressing the largest possible value in the verbalization of a multidigit number. For example, in English, the written string of Arabic digits *1200* can be verbalized as *twelve hundred*, as *a thousand two hundred*, or as *one thousand two hundred*. The verbalizations *a thousand two hundred* and *one thousand two hundred* are maximally fused because the word *thousand* has the largest value of a single English word that could be included in a group of words expressing 1200. The expression *twelve hundred* is not maximally fused because it does not include *thousand*, the English word with the largest value expressible in the number 1200.

Olsson goes on to explain that in some languages, year numerals are maximally fused. In Spanish, for example, there is no difference between the year numeral 1961 and the numeral used to count 1,961 objects: *mil noveciento sesenta y uno*. In English, however, while the counting number could be verbalized as either *one thousand nine hundred [and] sixty-one* or as *nineteen hundred [and] sixty-one*, the year numeral 1961 would never be verbalized with the words *thousand* or *hundred*. There is only one acceptable formulation in English, *nineteen sixty-one*. This lack of variation held true in English for the years through 1999 and the maximally-fused year of 2000. The year numeral 2001 brought some variety in that speakers could elect to insert the word *and* before expressing the final digit of the year numeral. The approach of the year 2010, and the need to refer to it in speech by its year numeral, created an option to express the year numeral in spoken English without maximal fusion. A variation in patterns of verbalization (the option of *twenty ten* or *two thousand [and] ten*) has become likely where for decades most speakers had only one option for the verbalization of a year numeral in English.

A frequent source of peer-reviewed research reports (e.g., Deviyani & Black, 2022; Pramanik and Hussain, 2019; Zhang et al, 2019) that refer to verbalization of year numerals is information systems science. One prominent sector of information science is the processing and synthesis of spoken human

language. Equipment must transcribe numbers with appropriate punctuation: is *two fifty* a time, 2:50, or a price, £2.50? It needs to recognize various verbalizations of the same string of digits (both *twelve hundred* and *one thousand two hundred* are 1200) and produce the appropriate verbalization of a multi-digit number (the U.S. telephone area code 312 is never *three hundred twelve,* always *three one two*). Research into speech synthesis and its application, text-to-speech, deals with the appropriate chunking of multidigit numbers according to the conventions and practices of various languages. Examples of languages' differences in chunking include the Japanese 万, an amount which in English requires two distinct words, *ten thousand*; and the German placement of the addend before the decade, as in *fünfundzwanzig* (literally *five and twenty*) for twenty-five. A typical issue in this field is how to build and program devices to verbalize year numerals distinctly from other 4-digit numerals that resemble year numerals, and might otherwise be mistaken for them.

There are non-research publications that predict, sometimes prescribe, and sometimes discourage, various English verbalizations of year numerals in the third millennium. Some of these are opinion pieces addressing an audience already highly proficient in English, and some are intended as language instruction. None of these cites any systematic count of how English-speakers verbalize year numerals.

Writing in *Science and Society*, a social science journal unrelated to language teaching, Laibman (2008) predicted that English-speakers would verbalize the year numeral 2010 as *twenty ten*, in contrast to the 10 previous years of *two thousands*. He based this in two speculated principles of speech, inertia (speakers continue to use established patterns of verbalization) and economy (speakers opt to use fewer syllables). *Twenty oh nine* could not predominate over *two thousand nine* because both verbalizations have the same number of syllables, but *twenty ten* will gain wider use, Laibman predicts, than *two thousand ten* because it has one fewer syllable. In *American Speech*, Carson (2010) ponders a rivalry between *twenty ten* and *two thousand ten* and cites broadcast style manuals. On December 31, 2009, a reporter for the CBS network in the U.S. "appears to correct his cohost and himself when they both say two thousand ten" (p. 354). In the realm of broadcast news, this researcher was quite surprised to hear Alex Haskell, a reporter for an NBC network affiliate in the American state of Maine (News Center Maine, 2025) verbalize the year 2007 as *twenty oh seven*.

But a post (rib, n.d.) in an online forum called *Pain in the English*, dedicated to "questioning, nitpicking, and debating the quirks and rules of the English language," claimed that veteran American newsman Charles Osgood, active on the CBS network from 1971 to 2017, encoded the years 2001 to 2009 as *twenty oh one* to *twenty oh nine* when he spoke. (The question of whether some year numeral verbalization patterns are the particular practices of broadcast journalists will be revisited in this

report.) In the brief but lively forum debate (1 post plus 12 comments) outlined in the post's title, there were 2 posts predicting both the greater use of *twenty ten* and less use of *two thousand ten*; 3 posts actively discouraging the use of *two thousand ten*; and 4 posts advocating in favor of *twenty ten*. Some participants tolerated the verbalization *two thousand ten*, but no one argued in its favor. On the *New Yorker* website, Hertzberg (2010) predicted and celebrated the end of the *two thousand* verbalization pattern. He also speculated that the pattern owed its wide use to the popularity of the 1968 movie 2001: A Space Odyssey. However, Ramirez (2010) notes in a *New York Times* column that the year numeral 2001 is never spoken in either this film or in the trailer that advertises it.

Materials aimed at English learners encouraged and discouraged a range of verbalization patterns for year numerals of the third millennium. A single source might encourage the use of more than one pattern. The pattern of *twenty ten, twenty eleven*, and so on was advocated for by Boyarskaya et al. (2012), Pirlo (2023), Puzatuykh (2018), Tanant (2014), and GrammarBook (2011). Pirlo (2023), Puzatuykh (2018), and GrammarBook (2011) advised the use of *two thousand ten*. Collishaw (2002), Puzatuykh (2018), and Tanant (2014) encouraged the verbalization pattern *two thousand and ten*. Interestingly, Walter (2017) claimed distinction in national dialects: *two thousand ten* supposedly leans American, while the British are said to favor *two thousand and ten*. This proposed international distinction is a field of inquiry for another piece of research. English instruction materials took active stances against only one pattern of year numeral verbalization, the use of *and* after *thousand*, as in *two thousand and ten*. GrammarBook (2011) discouraged the use of the word *and* in year numeral verbalization in any context, while Pirlo (2023) advised against its use in American dialects. (See Rice (1936) for arguments for and against the use of the word *and* in numeral verbalization.)

A YouTube video for English learners (English with Emma・engVid, 2024) explained how to say the year in English, and dealt with the possibility of variation for year numerals after 2009. The teacher in the video, Emma, is Canadian and therefore uses North American English, as do 59 of the 64 speakers included in this paper's corpus. Emma tells the viewer that for the year numeral 2010, both *two thousand and ten* and *twenty ten* are acceptable, but that *twenty ten* is more common because this verbalization is "faster" (English with Emma・engVid, 2024, 02:53). She never offers any evidence for this claim, which is understandable, given the context. (Research into this question, and the publicization of its results, would be of great benefit to situations like this.) She continues with this stance throughout the year numerals, always accepting the maximally fused verbalization starting with *two thousand*, but always recommending and modeling the verbalization that starts with *twenty*. She refers to the choice of verbalization pattern as "pronunciation," (0:05) which is, again, understandable in the circumstances of a free, publicly available video for learners of English. After the word *thousand*, she always includes the

word *and*, in both speaking and writing. YouTube offers an automatic transcript of the video, of interest to researchers who investigate machine processing of spoken language: “Do we pronounce this as 2005 or do we pronounce this as 2005 or 2005? Well, the way we actually pronounce this year is 2005.” (1:09) When the teacher speaks very slowly, automatic transcription verbalizes the year numerals: “We can also say ‘two thousand and ten,’ but in general, native speakers of English prefer to say ‘twenty ten’.” (2:42) This researcher wonders whether the transcription software produced the desired verbalization not only because of the deliberately reduced speed of the teacher’s voice but because alternatives to the maximally fused verbalization of 2005−of all year numerals from 2000 to 2009−are so rare that the software is not programmed to deal with them beyond writing them in digits.

This researcher’s use of the Google Scholar site failed to reveal any peer-reviewed report of an investigation of a spoken corpus in which verbalizations of any year numerals were identified and counted. Besides its hoped-for contribution to general description of English vocabulary use and verbalization of an extremely widespread type of numeral, a useful application of this research would be to tell English learners what to listen for in English dates, and the limits of what is considered appropriate spoken production.

## Method

This is a quantitative study, an identification and count of various ways English-users (mostly from the United States) verbalize the year numeral 2010 in the productive spoken domain of English. The researcher conducted this investigation because, as this study’s results will show, there is more than one acceptable verbalization of the year numeral 2010 in English, yet no study has been published comparing incidences of the different verbalizations.

The researcher identified utterances of the year numeral 2010 spoken in English on *This American Life*, a radio show produced in the United States in collaboration with Chicago Public Media and broadcast on many National Public Radio stations, also available online for streaming and download. *This American Life* is a weekly, mostly hour-long, general-interest program with themed episodes usually divided into “acts” like theatrical plays, with occasional focus on a single narrative. Spoken text produced by the show includes scripted monologues and commentary; the occasional multi-actor short drama; interviews; and speech and conversations ranging from extemporaneous to fully spontaneous; with examples of all these recorded in and out of the studio. *This American Life* has broadcast since 1995. As of this writing, more than 800 episodes, all named and numbered and indicating the date of first broadcast, are available on the show’s website, both as written transcripts and as audio recordings. With

rare exceptions (e.g., a child speaking to a reporter in public), these transcripts include each speaker's full name and enough context to access demographic information about that individual speaker.

The corpus includes all episodes available at the website www.thisamericanlife.com in both written transcript and audio format. If either format of an episode was not available at the website, the researcher did not include that episode in the corpus. This resulted in a corpus of 786 episodes, from the 2022 re-broadcast of Episode 109, "Notes on Camp," first broadcast August 28, 1998, to Episode 864, "Chicago Hope," first broadcast July 18, 2025, with a total of approximately 7,000,000 running words in spoken English.

The researcher used *This American Life* as a linguistic corpus of convenience. The availability of *This American Life'*s transcripts makes it easy to locate incidences of items of productive speech−in this case, any utterance of the year numeral 2010−within audio recordings, and then to listen for, identify, and annotate the verbalization pattern along with the episode number and name and its first date of broadcast; an approximate time marker within the episode's audio recording; the speaker's name; and whether that speaker used a dialect of North American English, or another dialect from outside the United States and Canada.

Each utterance of the year numeral 2010 was labeled alphabetically in the order in which it appeared in the broadcast series, using the letters A through FL. For example, the first entry, 2010A, appeared in Episode 237, "Regime Change," at approximately 31 minutes and 31 seconds within the episode. 2010EM and 2010EN both appear in Episode 641, "The Walls," spoken by the same person, a broadcast journalist named Lizzie Presser. As two utterances distinct from each other, they receive two separate labels. 2010EM precedes 2010EN alphabetically because it also precedes it in time, at approximately 56 minutes 32 seconds versus 1 hour 8 minutes 50 seconds.

The researcher marked each utterance of the year numeral 2010 for one of three verbalization patterns: $\alpha$ (Alpha) for *twenty ten*, $\beta$ (Beta) for *two thousand ten*, or $\gamma$ (Gamma) for *two thousand and ten*. The Greek alphabetical order does not indicate frequency or anyone's preference or theories related to grammatical form. *Twenty ten* received the $\alpha$ label because in *This American Life'*s first utterance of the year numeral 2010, the above-mentioned 2010A, first broadcast on April 18, 2003, the speaker, broadcast journalist Sarah Koenig, verbalized the year numeral 2010 as *twenty ten*. Pattern $\beta$, *two thousand ten*, was the next produced, in entry 2010B by host Ira Glass, in Episode 329, "Nice Work If You Can Get It," first broadcast April 6, 2007. Pattern $\gamma$, *two thousand and ten*, did not emerge until January 1, 2010 in Episode 397, titled "2010." The researcher chose to label verbalization patterns with

lower-case Greek letters for their ease of distinctiveness from the conventional Roman letters used in written English.

Below is a typical entry for a single utterance:

2010CO　β two thousand ten　525 Call For Help　09-May-14　06:25　Ira Glass　male　NAE

The year numeral 2010 was verbalized in utterance 2010CO as Pattern β (*two thousand ten*) in Episode 525, "Call For Help," first broadcast on May 9, 2014, at 6 minutes 25 seconds into the episode by host Ira Glass, a male user of North American English.

## Results

The corpus under investigation includes 786 episodes of *This American Life* available at the show's website, www.thisamericanlife.com, from August 28, 1998 to July 18, 2025. The researcher found 168 utterances of the year numeral 2010 and labeled them alphabetically from 2010A to 2010FL. The 168 utterances were found in 91 separate episodes and were spoken by 64 separate individuals. Males spoke 120 of the 168 utterances, while females and one speaker not on the gender binary spoke 48 utterances out of the total 168.

Listeners should perhaps expect that on a program titled *This American Life* 160, or 95.24%, of the 168 utterances were spoken in some variety of North American English. (Speakers who did not use North American English were from Israel, China, Mexico, and the United Kingdom. Russian-born Masha Gessen was counted among users of North American English due to noticeable North American features in Gessen's spoken English, including vocabulary−e.g., *college* instead of *university*; *vacation* instead of *holiday*−and pronunciation, especially North American vowels and the syllable-final rhotic /r/.)

As shown in Table 1, the most common verbalization pattern used for the year numeral 2010 was Pattern β , *two thousand ten,* spoken 102 times, followed by Pattern α , *twenty ten*, spoken 50 times, and Pattern γ , *two thousand and ten,* spoken only 16 times. Beyond these 3 patterns, no other verbalization pattern was found for the year numeral 2010.

Table 1　*Verbalization patterns for year numeral 2010, all speakers*

| | male | female and non-binary | Pattern total |
|---|---|---|---|
| Pattern α : *twenty ten* | 31 | 19 | 50 |
| Pattern β : *two thousand ten* | 79 | 23 | 102 |

| | | | |
|---|---|---|---|
| Pattern γ : *two thousand and ten* | 10 | 6 | 16 |
| | Male total= 120 | Female and Non-binary total= 48 | Grand total= 168 |

## Individual Speakers of Multiple Patterns

Of 64 individual speakers of the year numeral 2010, 42 individuals produced it only once, but 22 spoke it more than once (Table 2). Of these 22 speakers, 16 always produced the same verbalization pattern when they spoke the year numeral 2010. The remaining 6 speakers, in discrete utterances, used different verbalization patterns to speak the year numeral 2010. However, not one speaker—not even Ira Glass, who produced more utterances of the year numeral 2010 than any other speaker, and whose production this paper will more closely investigate later—used all three verbalization patterns.

Table 2 *Speakers of multiple utterances of the year numeral 2010*

| | Pattern α (*twenty ten*) | Pattern β (*two thousand ten*) | Pattern γ (*two thousand and ten*) |
|---|---|---|---|
| Shalom Auslander | 0 | 2 | 0 |
| Sara Blaisdell | 0 | 2 | 0 |
| Alex Blumberg | 0 | 3 | 0 |
| Ben Calhoun | 7 | 2 | 0 |
| Dana Chivvis | 2 | 0 | 0 |
| Sean Cole | 0 | 6 | 0 |
| Ira Glass | 14 | 39 | 0 |
| Jack Hitt | 0 | 0 | 4 |
| Joe Hubbard | 2 | 0 | 0 |
| Chana Joffe-Walt | 0 | 5 | 0 |
| Etgar Keret | 0 | 1 | 3 |
| David Kestenbaum | 0 | 2 | 0 |
| Sarah Koenig | 6 | 0 | 0 |
| Frank Langfitt | 0 | 0 | 3 |
| Joa Marcu | 0 | 2 | 0 |
| Miki Meek | 0 | 3 | 0 |
| Jonathan Menjivar | 0 | 3 | 0 |

| | | | |
|---|---|---|---|
| Lizzie Presser | 1 | 0 | 1 |
| Nadia Reiman | 1 | 1 | 0 |
| Christopher Rhoads | 2 | 0 | 0 |
| Rob Schmitz | 0 | 4 | 0 |
| Nancy Updike | 1 | 4 | 0 |

These multiple utterances from individual speakers show the same prevalence of patterns as the overall count: Pattern β was used the most, 79 times; Pattern α came in second at 36 utterances; and Pattern γ was used the least often, 11 times.

## Analysis by Speaker Gender

Thirty-five males (34 men and 1 boy) spoke the year numeral 2010 a total of 120 times, and most often verbalized it as Pattern β : *two thousand ten,* as seen in Table 3, below.

Table 3 *Verbalization of year numeral 2010 by male speakers*

| | male raw count | % of male total (120) | % of grand total (168) |
|---|---|---|---|
| Pattern α : *twenty ten* | 31 | 25.83% | 18.45% |
| Pattern β : *two thousand ten* | 79 | 65.83% | 47.02% |
| Pattern γ : *two thousand and ten* | 10 | 8.33% | 5.95% |

The male who produced the most utterances of the year numeral 2010 was program host Ira Glass, who was also the top producer of any gender. Ira Glass spoke the year numeral 2010 53 times: 31.55% of the total 168, and 44.16% of all 120 male utterances of the year numeral 2010.

The researcher divided all speakers into two gender groups: one male group, and one female and non-binary group. The one non-binary speaker is journalist Masha Gessen. Gessen was assigned female at birth in 1967 and publicly announced the use of they/them personal pronouns in 2020. Gessen mentioned in a 2020 Russian-language interview (Shainyan, 2020) that they use grammatically feminine past tense verbal suffixes when speaking their native Russian. It was these two indicators, gender assignment at birth and grammatically feminine language choices, which convinced the researcher to include the non-binary Gessen in the same group with women.

Another group member of interest is a British zookeeper named Melissa Bushell who, knowing that her voice was to be recorded and broadcast in a segment of Episode 776, "I Work Better on Deadline," asked that the audience be told that she had "recently transitioned" to female (David Kestenbaum reporting in Glass, 2022, 52:08). This researcher chooses to honor Bushell's preference and include her among female and non-binary speakers of the year numeral 2010.

A group of 29 (28 females and one non-binary person) spoke the year numeral 2010 a total of 48 times. The top female/non-binary producer of utterances of the year numeral 2010 was broadcast journalist Sarah Koenig, who spoke it only 6 times.

Table 4 *Verbalization of year numeral 2010 by female and non-binary speakers*

| | female and non-binary | % of female/ non-binary total (48) | % of grand total (168) |
|---|---|---|---|
| Pattern α : *twenty ten* | 19 | 39.58% | 11.31% |
| Pattern β : *two thousand ten* | 23 | 47.92% | 13.69% |
| Pattern γ : *two thousand and ten* | 6 | 12.5% | 3.57% |

As seen in Table 4 above, female/non-binary utterances of the year numeral 2010 were more evenly split between Patterns α and β, with only an 8.34% difference between the two choices, while the male preference for Pattern β over Pattern α was a much stronger 40.83%. Neither group favored Pattern γ, but female/non-binary people were 4.17% more likely to use it.

## Comparison of Host Ira Glass with Other Speakers

Since program host Ira Glass accounted for 31.55% of all utterances of the year numeral 2010, the researcher was interested to see whether his verbalization patterns differed from those of other speakers.

Table 5 *Verbalization of year numeral 2010 by host Ira Glass compared with all other speakers*

| | Ira Glass raw count | % of Ira Glass total (53) | Non-Ira Glass raw count | % of non-Ira Glass total (115) |
|---|---|---|---|---|
| Pattern α : *twenty ten* (52) | 14 | 26.41% | 36 | 31.3% |
| Pattern β : *two thousand ten* (104) | 39 | 73.58% | 63 | 54.78% |
| Pattern γ : *two thousand and ten* (16) | 0 | 0% | 16 | 13.91% |

The most striking difference between Ira Glass and all other speakers was his complete eschewal of Pattern γ in the face of other speakers' choice of this pattern in 13.91% of their utterances of the year numeral 2010. Another noticeable feature was his stronger preference for Pattern β over Pattern α, a difference of 47.17% versus all other speakers' weaker preference of only 23.48%.

With contrasts in mind between the male group and the female/non-binary group, the researcher compared Ira Glass's results with these two groups separately.

Table 6 *Verbalization of year numeral 2010 by host Ira Glass compared with female and non-binary speakers*

| | raw count from Ira Glass | % of Ira Glass total (53) | raw count from female and non-binary | % of female and non-binary total (48) |
|---|---|---|---|---|
| Pattern α : *twenty ten* | 14 | 26.41% | 19 | 39.58% |
| Pattern β : *two thousand ten* | 39 | 73.58% | 23 | 47.92% |
| Pattern γ : *two thousand and ten* | 0 | 0% | 6 | 12.5% |

Table 7 *Verbalization of year numeral 2010 by host Ira Glass compared with other male speakers*

| | raw count from Ira Glass | % of Ira Glass total (53) | raw count from non-Ira Glass males | % of non-Ira Glass male total (67) |
|---|---|---|---|---|
| Pattern α : *twenty ten* | 14 | 26.41% | 17 | 25.37% |
| Pattern β : *two thousand ten* | 39 | 73.58% | 40 | 59.7% |
| Pattern γ : *two thousand and ten* | 0 | 0% | 10 | 14.93% |

As seen in Table 6 above, like other male speakers, Ira Glass differed from speakers from the women and non-binary group in producing Pattern γ much less often (that is, not producing it at all), and in his stronger preference for Pattern β.

Table 7, above, shows that Ira Glass's verbalization patterns for the year numeral 2010 would be quite similar to those of other males were it not for his complete avoidance of Pattern γ.

## Discussion

The purpose of this study was to identify and count verbalization patterns in English speech for the year numeral 2010. The researcher selected 2010 because it is the new millennium's first year numeral without a 0 in the third digit position. The application of these findings will be of particular use to

teachers of English as a second or foreign language: learners need to know that there's more than one acceptable way to say the year "2010" in English, and they need to know which ones they are more likely or less likely to hear.

Repeatedly searching the Google Scholar search engine with permutations of such terms as *year, English, verbalization, corpus, nineteen, twenty, oh*, and *thousand*, the researcher was unable to locate any peer-reviewed quantitative study on patterns of spoken verbalization of any year numeral in English.

## Limitations

This study is an investigation of only 168 utterances from a small corpus of about 7,000,000 running words in spoken English. A larger number of utterances of the year numeral 2010 could be found in a larger corpus of spoken English. However, the convenience of this small corpus, accessible at no cost, and easy to analyze even with very simple, familiar technology, appealed to the researcher.

This population of speakers of the year numeral 2010 is not very diverse. Regardless of who is interviewed, year numerals on *This American Life* tend to be spoken in narrative summaries by the shows' presenters, who tend to be educated White Americans (or at least speakers of American accents) born between 1950 and 1990. The researcher had hoped to compare the utterances of speakers of North American dialects with those of speakers of other dialects, but of 168 utterances of the year numeral 2010, only 8 came from non-North American dialects, spoken by only 5 individuals.

Additional variables were sought among the population of speakers. Speaker gender was not difficult to investigate, being easily confirmed by third-person pronoun references to speakers in audio recordings and online profiles. With such a large portion of utterances produced by American-accented speakers, the researcher had considered a comparison of utterances made by different ethnic populations within the United States. For example, do Black speakers tend to use a different verbalization pattern for the year numeral 2010 than non-Black speakers? However, the confirmation of ethnic information beyond nationality proved difficult to confirm, and research into individual identities seemed both impractical and invasive.

The researcher had initially been struck by a perception of a greater use of Pattern $\gamma$, *two thousand and ten*, by less educated speakers and had hoped to investigate this possibility. But education levels could not be known for non-famous speakers, and the variable was too difficult to define.

## Future Research

The quantification of verbalization patterns of the year numeral 2010 from a much larger corpus, or from a corpus with much more demographic information available about the speakers, including ethnicity distinct from nationality, age, education level, and greater specificity of regional dialects, would provide much more detail of how English-speakers say the year numeral 2010.

This researcher is especially interested in a comparison of how broadcast journalists verbalize third-millennium year numerals, with the practices of the general population. As stated above in the Limitations section, I thought I noticed a greater frequency of the use of Pattern γ , *two thousand and ten*, by less educated speakers. This led me to look at a broader set of data including all spoken year numerals, not just 2010, after 2009 in the *This American Life* corpus. I looked for verbalization patterns used by speakers whom I knew to be university-educated: journalists. I then began to wonder whether the difference is based not so much in education, but in specific profession. In America, do broadcast journalists verbalize third-millennium year numerals differently from the general population? (Recall that host Ira Glass never used Pattern γ .) I considered annotating speakers in the corpus as either broadcast journalists or general population, but found that among the interviewees were professional broadcast journalists who were not working for *This American Life*. When a broadcast journalist is an interviewee, and not an interviewer, does he cease to be a broadcast journalist? I also found reporting done by people who worked for *This American Life*, but usually in other roles. When a writer or sound editor reports a story, does she become a reporter? There were also guest reporters whose usual professions were in comedy or theater. Should we call them broadcast journalists when they read current-affair scripts over the radio? So I made the decision not to single out broadcast journalists and compare them with the general population, but was intrigued by 3 incidents noted in this paper's literature review. Carson (2010) described two reporters making a deliberate change from *two thousand ten* to *twenty ten*. American broadcast journalist Charles Osgood started his verbalizations of third-millennial years with *twenty* instead of *two thousand*, even for years with a third digit 0 (Corliss, 2004; Kyff, 2009; Ramirez, 2001; rib, n.d.). In 2025, an NBC reporter in Maine verbalized the year numeral 2007 as *twenty oh seven* (News Center Maine, 2025). Is this a practice more common among broadcast journalists, and is it due to an effort to save time within a broadcast? The former question might be answered by investigation of a much larger corpus that includes many more year numeral verbalizations from non-broadcasters for comparison.

Considering the history of changes in how anglophones have verbalized first-millennium 4-digit year numerals with a third digit 0, it would be useful to track changes over time in verbalizations of year

numerals after 2009, that is, from 2010 onward. Researchers might try to predict how English-users will verbalize the year numeral 2010 in, for example, the year 2035. This information would be especially useful to teachers of English as a second or foreign language as they prepare their students to be able to listen for year numerals in spoken English.

## References


Boyarskaya, A., Ladutko, N., & Mitkovets, T. (2012). *Spoken English for Transportation*. Belarusian National Technical University. https://rep.bntu.by/handle/data/5923

Carson, C. (2010). Among the new words. *American Speech, 85*(3), 352-365. https://doi 10.1215/00031283-2010-020

Collishaw, B. (2002). FAQs on Writing the Date. *Terminology Update 35*(2), 12. https://www.noslangues-ourlanguages.gc.ca/fr/favourite-articles/faqs-on-writing-the-date

Corliss, R. (2004). That old feeling: Sunday morning going strong. *Time Magazine*. https://time.com/archive/6912366/that-old-feeling-sunday-morning-going-strong/

Deviyani, A., & Black, A. W. (2022). Text normalization for speech systems for all languages. In *Proceedings of the First Workshop on Speech for Social Good (S4SG)*, 20-25. https://doi.org/10.21437/s4sg.2022-5

English with Emma・engVid (2024, December 30). *How to Say Years in English* [Video]. YouTube. https://www.youtube.com/watch?v=PN8P8F-Pgtc

Glass, I. (Host). (2022, July 29). I work better on deadline. (No. 776) [Audio podcast episode]. In *This American Life*. WBEZ Chicago. https://www.thisamericanlife.org/776/i-work-better-on-deadline

GrammarBook. (2011, April 29). Yes, you are correct. The word “and” should not be included (one hundred fifty). The year 2011 may be written [Comment on the blog post “Writing numbers as both numerals and words”]. GrammarBook. https://www.grammarbook.com/blog/numbers/writing-numbers-as-words-2/#comment-5779

Hertzberg, H. (2010, January 1). Twenty-Something. *News & Politics* [Newsletter]. https://www.newyorker.com/news/hendrik-hertzberg/twenty-something-2

Kyff, R. (2009, June 5). How to pronounce 2009? Let us count the ways. *Hartford Courant*. https://www.courant.com/2009/06/05/how-to-pronounce-2009-let-us-count-the-ways/

Laibman (2008). Year naming, millennial transitions, and English: A trivial observation, and prediction. *Science and Society, 72*(4), 383-388.

News Center Maine (2025, October 22). *Senate candidate Graham Platner bares his chest amid tattoo scandal* [Video]. News Center Maine. https://www.newscentermaine.com/video/news/politics/maine-politics/senate-candidate-graham-platner-bares-his-chest-amid-tattoo-scandal/97-e2cc705c-35a3-42f2-acfe-ea6b109ffd9d

Olsson, M. (1988). How dates are made. *Working papers/Lund University, Department of Linguistics and*

*Phonetics*, *33*, 173-191.

Pirlo, F. (2023, April 13). Pronouncing Years in American English [Audio podcast episode]. In *Everyday Grammar*. Voice of America News. https://learningenglish.voanews.com/a/pronouncing-years-in-american-english/7045997.html

Pramanik, S., & Hussain, A. (2019). Text normalization using memory augmented neural networks. *Speech Communication*, *109*, 15-23. https://doi.org/10.48550/arXiv.1806.00044

Puzatuykh, A. (2018). *English for Bachelor's Degree Students*. Bunin Yelets State University.

Ramirez, A. (2001, December 30). Twenty-oh-brother. *The New York Times*. https://www.nytimes.com/2001/12/30/weekinreview/twentyohbrother.html?searchResultPosition=1

rib (n.d.). Twenty-ten vs Two thousand-ten. *Pain in the English*. https://painintheenglish.com/case/4273

Rice, W. (1936). Usage counsel: Varia. *The English Journal, 25*(1), 67-69.

Shainyan, K. (2020, February 18). Маша Гессен: о Трампе, тестостероне и терроре [Video]. (Masha Gessen: on Trump, testosterone, and terror.) YouTube. https://www.youtube.com/watch?v=BAYNKHP-kVQ

Tanant, K. (2014). *Tax Management – L'Anglais de la fiscalité*. Ophrys.

Walter, L. (2017, October 4). 1066 and all that: How to say years. *About Words*. https://dictionaryblog.cambridge.org/2017/10/04/1066-and-all-that-how-to-say-years/

Zhang, H., Sproat, R., Ng, A. H., Stahlberg, F., Peng, X., Gorman, K., & Roark, B. (2019). Neural models of text normalization for speech applications. *Computational Linguistics, 45*(2), 293-337.