

サンプルID	データ公開時期	修正前	修正内容
IID13-ST1.txt	第一次	データ全体の調査者(K)と学習者(K)の発話者記号が入れ替わっている。	データ全体の調査者(K)と学習者(K)の発話者記号をすべて入れ替えた。
SES04_I	第一次	発話番号03200行に「? 過剰使用」という形態素解析処理のための情報が出ていた。 03200-K 今はそんなに若くないだ? 過剰使用と思いますが、友達	「若くないだ」の後の「? 過剰使用」を削除した。 03200-K 今はそんなに若くないだと思いますが、友達
CCM14_I	第一次	02480行から02530まで話者記号が入れ替わっている 02470-C [あー=F]そうですか 02480-K 小学校とか中学校とか高校で、好きだった先生はいますか? 02490-C えーいます、えー小学校の先生 02500-K うーどんな先生でした? 02510-C えーとても、[う、ん、うー、え先生は、彼女、です 02520-K 女の先生	入れ替わっていた話者記号を修正した。 02470 C あーそうですか、小学校とか中学校とか高校で、好きだった先生はいますか? 02480-K えーいます、えー小学校の先生 02490-C うーどんな先生でした? 02500-K えーとても、う、ん、うー、え先生は、彼女、です 02510-C 女の先生 02520-K ん、はい
JJC35_RP1	第一次	個人情報を伏字化した	【人名1】とした（音声も編集済み）
JJC46_RP1	第一次	個人情報を伏字化した	【人名1】とした（音声も編集済み）
CCM14-I	第一次	赤字の「その」がXタグになっていた。 03610-K そのより、えーいなが(田舎)が、うんちょっと、寂しい(さびしい)の感	赤字の品詞を解析困難箇所から連体詞に変更した。
CCM18-I	第一次	Fタグの「F」が本文に出ていた 00300-K (末尾)コンピューターをあーF堪能しています	本文に出ていたFタグの「F」を削除 0300-K (末尾)コンピューターをあー堪能しています
CCM22-I	第二次	音声ではピー音で伏せられている個人情報の一部が、【】でくられていなかった。 01380-K あのゆ、あのうちはーえーと【公園の名前1】公園という公園の隣で 01400-K で、あの、【公園の名前1】公園に行ったら、絶対中に入らないでください	個人情報の【】外の「公園」を削除。 01380-K あのゆ、あのうちはーえーと【公園の名前1】という公園の隣で 01400-K で、あの、【公園の名前1】に行ったら、絶対中に入らないでください
CCM30-I	第二次	音声ではピー音で伏せられている個人情報の一部が、【】でくられていない。 00440-K そして(はい)日本語【学科名1】を選んでちょうどえと【大学名1】大学の【学科名1】に入りました	個人情報の【】外の「大学」を削除。 00440-K そして(はい)日本語【学科名1】を選んでちょうどえと【大学名1】の【学科名1】に入りました
CCM52-I	第一次	赤字の箇所の品詞が「名詞-固有名詞」になっていた。 0290-C なんとゆう	赤字を名詞-固有名詞から動詞-一般に変更した。
EAU30-RP2	第二次	「では」の前に文字化漏れがあった。 00030-C ではロールプレイ2を始めます	漏れていた「はい」を追加した。 00030-C はい、ではロールプレイ2を始めます
EAU37-I	第一次	Fタグの「F」が本文に出ていた 01920-K あーFたぶん都会の生活?	本文に出ていたFタグの「F」を削除 01920-K あーたぶん都会の生活?

EAU40-I	第一次	赤字の部分の品詞が誤解析になっていた。 00050-K はい車では <b>だいたぶん</b> 一時間かかりますけどー、交通 ああバスを使うと二時間かかります	誤解析になっていた赤字の部分を2短単位に分割し、それぞれ以下 の品詞を付与した。 だい: 解析困難箇所 たぶん: 副詞
EAU40-I	第一次	タグの不備により、赤字の不要部分が本文に出ていた。 00340-C 作れるんです、か <b>作れるんですか</b> ?	赤字の箇所を削除した。
EUS24-RP1	第二次	文字化作業時のメモが残っていた。 00030-K { <b>ノックの音?非言語行動の表示</b> でよいか?}	不要なメモを削除した。 00030-K {ノック音}
EUS24-RP2	第二次	文字化作業時のメモが残っていた。 00040-K { <b>ノックの音?非言語行動の表示</b> でよいか?}	不要なメモを削除した。 00040-K {ノック音}
EUS49-I	第二次	発話にないものが文字化されていた。 03670-C はい、じゃ、えーと、これから十年か、二十年、後(あ と)で、住むとしたら、都会がいいですか、田舎がいいですか? <b>答 え</b>	赤字部分を削除した。
FFR07-D	第一次	Gタグの付与ミスにより検索システムでは以下の文字列で検索さ れるようになっていた。 0300-K それがー、ん生徒達とか学生だ、が、がー、んー、 <b>乗っ ています</b>	Gタグ修正の範囲ではないため、Gタグを削除し、発話通り「乗れ ています」で形態素解析し、中納言に搭載した。 0300-K それがー、ん生徒達とか学生だ、が、がー、んー、 <b>乗れ ています</b>
FFR07-I	第一次	Fタグの「F」が本文に出ていた 01120-K はー <b>F</b> 残念ですけど{笑}	本文に出ていたFタグの「F」を削除 01120-K はー残念ですけど{笑}
FFR10-D	第一次	Fタグの「F」が本文に出ていた 00100-K 家のーえーとー坂ーで? えー <b>F</b> 、{笑}殺させた人が、 います?	本文に出ていたFタグの「F」を削除 00100-K 家のーえーとー坂ーで? えー、{笑}殺させた人が、い ます?
FFR27-I	第一次	以下のFRの部分分割されて解析されていた。 00610-C えっとー <b>FR</b> さん、のー出身ー地、故郷(こきょう)ふるさ とはどこですか?	FRの部分を1短単位として修正した。
FFR50-I	第一次	Fタグの「F」と不要な記号が本文に出ていた 03070-C あー <b>F</b> やっぱりインターネット使うんですね	本文に出ていたFタグの「F」を削除 03070-C あーやっぱりインターネット使うんですね
IID27-I	第一次	赤字の箇所の品詞が「助詞-終助詞」になっていた。 0180-K 乗り物、 <b>やー</b>	赤字部分の品詞を助詞-終助詞から助詞-副助詞に変更した。
IID27-I	第一次	赤字の箇所の品詞が「感動詞ー一般」になっていた。 1910-C 弟さん、 <b>は?</b>	赤字を感動詞ー一般から助詞-係助詞に変更した。
IID46-I	第一次	Fタグの「F」が本文に出ていた 01190-K んー <b>F</b> はい、すち(土)で	本文に出ていたFタグの「F」を削除 01190-K んーはい、すち(土)で
JJC11-I	第一次	Fタグの「F」が本文に出ていた 00310-C えー <b>F</b> 、けっこう遠いんですねー	本文に出ていたFタグの「F」を削除 00310-C えー、けっこう遠いんですねー

JJJ02-I	第二次	漢字変換のミスがあった。 02340-C あのーうん、あのす、ちょっと桜の、あの塩漬けて言うか(ゆうか)、<うんうんうん>桜の、が <b>載ったー</b> 、あのマドレーヌとか、<へー>そうゆうのを出してるお店もありますね	赤字の箇所を修正した。 02340-C あのーうん、あのす、ちょっと桜の、あの塩漬けて言うか(ゆうか)、<うんうんうん>桜の、が <b>乗ったー</b> 、あのマドレーヌとか、<へー>そうゆうのを出してるお店もありますね
JJJ11-FS	第一次	FSの年齢の欄が間違っていた。	FSの年齢を20歳に修正した。
JJJ23_D	第二次	漢字変換のミスがあった。 00010-C はい、ではIDを <b>行</b> ってください	「行」を「言」に修正した。 00010-C はい、ではIDを <b>言</b> ってください
JJN27-I	第二次	「ブブー」がXタグになっていた。 06630-C うんすぐにみんな <b>ブブー</b> って	解析困難箇所を副詞にした。 06630-C うんすぐにみんなブブー(擬音語・擬態語)ーって
KKD06-I	第一次	赤字「すじ」がXタグになっていた。 KKD06-I-04510-C 良かったですねー妹さん、隣、お婆さんの <b>すじ(/家の可能性あり)</b> にいたんですね	XタグをGタグにし、補足情報を削除。 04510-C 良かったですねー妹さん、隣、お婆さんの <b>すじ(家)</b> にいたんですね
KKD06-I	第二次	赤字の「いるー」が解析除外になっていた。 04700-K この(連体詞)インタビューの前も<うん>、ちょっと学校で<はい>ヒューマンブックと<はい>、 <b>いるー(/いう)</b> ちょっと、プロジェクトかなー<はい>イベントかながあったら<はい>、今、放送作家の働くーのす先輩がきて(来て)<はい>、ちょっと一時間ぐらい話しました	赤字の「いるー」をXタグからGタグで「いうー」に変更した。
KKD06-I	第一次	タグのミスにより、学習者の発話が修正されてしまっていた。 05120-K あ、あ、後(あと)で考えることはー<はい>、かり、たい(「変えたい」の誤用か)、あい、か <b>借りる</b> (/借りる/変える)かーもしれませんけどー<はい>、始める時はだめっと	学習者の発話通りに修正し、品詞は解析困難箇所とした。 05120-K あ、あ、後(あと)で考えることはー<はい>、かり、たい(「変えたい」の誤用か)、あい、 <b>かりーる</b> (/借りる/変える)かーもしれませんけどー<はい>、始める時はだめっと(だめと)かりーる
KKD10-I	第一次	赤字の補足情報を修正した。 01440-K 息子ーのことに對してなんか、面倒<うんうんうんうん>くれ、てな、なくて( <b>くれてなくての誤用か</b> )てなくて?<うん>いる[笑]じよ、き、じ、状況なんで、状況でしたんですけどー	補足情報を赤字のように修正した。 01440-K 息子ーのことに對してなんか、面倒<うんうんうんうん>くれ、てな、なくて( <b>くれてなくて</b> )?<うん>いる[笑]じよ、き、じ、状況なんで、状況でしたんですけどー
KKD10-I	第一次	赤字の通り、文字化にミスがあった。 03700-K はい危ないです、でも、でも、その(連体詞)頃(ころ)は寒すぎなんですけど<はい>人(ひと)達は、そこが見たくて<はい>結構 <b>行きそう</b> です	余分な「そう」とGタグを削除し、発話通りの文字列「行きそう」で解析し直した。 03700-K はい危ないです、でも、でも、その(連体詞)頃(ころ)は寒すぎなんですけど<はい>人(ひと)達は、そこが見たくて<はい>結構 <b>行きそう</b> です
KKD20-I	第一次	「まあ」の部分がYタグになっていた 00550-C へーわかりました、ふーん、じゃああの日本語、 <b>まあ</b> 交流のために・・・	赤字の「まあ」をYタグからFタグに変更した。
KKD34-ST2	第一次	赤字の箇所がGタグで動詞に修正されていた。 00090-K でん、そーれー、そして、ケンの夫婦が、おー窓ー、から、見て?、大丈夫だと言って(ゆって)?けん、かーん、警官さんと、 <b>あいい(/言い)</b> 、あ終わりました[笑]	赤字の品詞を動詞から解析困難箇所にした。 00090-K でん、そーれー、そして、ケンの夫婦が、おー窓ー、から、見て?、大丈夫だと言って(ゆって)?けん、かーん、警官さんと、 <b>あいい</b> 、あ終わりました[笑]

KKD45-RP2	第一次	赤字の箇所の品詞が形容詞になっていた。 0290-C あの人は <del>ねー</del>	赤字部分の品詞を形容詞から助詞に変更した。
RRS30-I	第一次	赤字の品詞が助詞だった。 02040-K <del>も</del> と思いますけど、お	赤字の品詞を助詞から解析困難箇所にした
SES28_RP1	第一次	タグの不備による文字化漏れがあった。 00060-K てんちょ(さん)	本来の発話と補足情報を修正した。 00060-K てんちょ(店長)さん
TTH03-I	第一次	読点が重なっている。 05140-K 休日にね、 <del>、</del> 休日になったら	読点が重なっていたところを削除。 05140-K 休日にね、休日になったら
TTR33-I	第一次	補足情報の付与位置にミスがある。 <プレインテキスト> 03120-K あー、ん、私は、その、車の、何(なん)ですか、あ、 私は、 <del>しゅ</del> 学校(/ <del>小</del> 学校/ <del>中</del> 学校)時(はい) <I-JAS中納言> 03120-K あー、ん、私は、その、車の、何(なん)ですか、あ、 私は、 <del>の</del> (/ <del>小</del> 学校/ <del>中</del> 学校)時(はい)	補足情報の付与位置にミスがあったため以下のように修正した。  03120-K あー、ん、私は、その、車の、何(なん)ですか、あ、 私は、 <del>しゅ</del> (/ <del>小</del> / <del>中</del> )学校の時(はい)、
VVN18-I	第一次	「はー」に不要な補足情報があった。 03780-K あーい、本、 <del>はー</del> (助詞ではない)本を読むこととか	不要な情報を削除し、品詞を感動詞-一般から解析困難箇所に修正した。 03780-K あーい、本、 <del>はー</del> 本を読むこととか
VVN26-RP2	第一次	タグのNが出ている。 00260-K あーそれー以外は<うん>あークックNのことを勉強してもいいですか？	本文に出ていたNタグの「N」を削除 00260-K あーそれー以外は<うん>あークックのことを勉強してもいいですか？