

منبع 1:

فایل های صوتی باید دارای این خصوصیات باشند:

1- بدون خطا:

هر گونه تپق یا خطا در گفتار فایل ها باعث کاهش بازده میشود. در صورتی که یک جمله دارای اشکال است یا تلفظ آن مشخص نیست یا به هر دلیلی در ضبط آن با مشکل مواجه هستید، از روی آن رد شده و جای آن فایل را خالی بگذارید. حتی اگر فقط نیمی از فایل ها با کیفیت خوب ضبط شوند بازده کار از داشتن همه فایل ها با کیفیت پایین بهتر است.

توجه داشته باشید که قسمت بندی جملات و پردازش های انجام شده روی متن توسط ماشین انجام شده و احتمال خطا در متن وجود دارد. در صورت مواجهه با خطا صرفا از روی آن عبور کنید.

2- نداشتن نویز زمینه:

بازده کار به نویز زمینه به شدت حساس است. تا حد امکان از نویز در فایل جلوگیری شود.

3- لحن ثابت:

جملات داخل فایل از منابع مختلفی تشکیل شده اند و در حالت عادی برای اجرای صوتی خوب این متن ها، لحن های متفاوت برای قسمت های مختلف و شخصیت های مختلف یک نکته مثبت به شمار می رود اما برای این کار بخصوص به هیچ عنوان نباید از این تکنیک استفاده کرد. هر خط از متن باید به صورت مستقل و بدون در نظر گرفتن خطوط قبل و بعد اجرا شود. لحن صحبت فارغ از موضوع و این که نقل قول از چه کسی است، باید با لحن ثابت و شخصیت ثابت اجرا شود به طوری که اگر یک جمله در دو قسمت از فایل تکرار شود، صدای خروجی باید بیشترین شباهت ممکن به یکدیگر را داشته باشد.

4- صدای طبیعی:

علاوه بر نکته قبلی باید به این نکته هم توجه داشت که اجرای صوتی نباید خشک و با لحن ماشینی باشد. هر چقدر اجرا و صدا طبیعی تر باشد، خروجی بهتر خواهد بود. مطلوب ترین سرعت گفتار، سرعت طبیعی است. نه به طور آزار دهنده ای کند و نه به صورت غیر قابل فهمی تند. در مورد تن صدا و گام هم به همین منوال. در ضمن به علایم نگارشی در جمله هم حتما توجه لازم را داشته باشید. جملات سوالی یا ویرگول ها در جمله حتما با دقت لازم اجرا شوند. در کل مهمترین خصوصیت اجرا طبیعی بودن و واضح بودن در عین ثبات در طول تمام فایل ها است.

توجه داشته باشید که تمام خطوط ارزش یکسانی دارند و پس از اجرا فایل ها به صورت رندوم کنار یکدیگر قرار خواهند گرفت. پس تمام تلاش خود را بکنید که پس از عوض شدن ترتیب هم احساس یکدستی صوت ها حفظ شود.

منبع 2:

خلاصه ای از نکات برای ضبط دیتاست برای متن به صوت:

مهم ترین نکته در این کار ثبات و همخوانی است. تمام صوت ها باید به صورتی باشند که انگار در روز و در یک اتاق ضبط شده اند.

یکی از روش های کار آمد برای رسیدن به این کیفیت، گفتن یک متن ثابت در اول هر قسمت و مقایسه آن با قسمت های قبلی برای رسیدن به جزییات مشابه است.

صوت های ضبط شده باید از نظر احساسی خنثی باشند. مثلا جملات غمگین نباید با لحن غمگین خوانده شوند. در طول ضبط گوینده باید یک "شخصیت" ثابت داشته باشد و تمام جملات باید متناسب با آن شخصیت خوانده شود.

مثلا ممکن است شخصیتی "سرزننده" انتخاب شود. پس در تمام طول ضبط باید این سرزندگی و خوشبینی حفظ شود و ثابت بماند. به خاطر رسیدن به ثبات هرچه شخصیت صوتی به حالت عادی گوینده نزدیک تر باشد نتیجه بهتر خواهد بود.

معمولا استودیو ها در هر فایل حدود 10 تا 24 جمله را ضبط میکنند. در این صورت نوشتن شماره فایل جلوی هر عبارت کافیت. در صورتی که تعداد بیشتری عبارت در هر فایل ضبط میشود نوشتن زمان هر عبارت در طول فایل هم لازم است.

فایل ها باید در فرمت wav یا aiff با کیفیت 44.1 kHz 16-bit یا بهتر باشند. کیفیت 48 kHz 24-bit حالت مناسب است و کیفیت های بالا تر مانند 96 kHz نیاز نیست.

کلمات باید دقیقا به صورتی که نوشته شده خوانده شوند. مثلا اگر کلمه "خواندند" در عبارت آمده بود نباید به صورت "خوندن" تلفظ بشود اما اگر در متن کلمه به صورت عامیانه وجود داشت، مثلا "اگه" (به جای "اگر") باید به همان صورت عامیانه، "اگه" خوانده شود.

برای کار ما فقط یک کانال صوتی کافی است و نیازی به ضبط استریو نیست. البته ضبط استریو میتواند برای حذف نویز استفاده شود اما فایل نهایی باید به صورت تک کاناله باشد.

پس از ضبط در صورت امکان و در دسترس بودن متخصص ادیت این این دو عملیات را روی فایل ها انجام دهید.

بعد از ضبط باید فایل ها را به فایل های یک عبارتی تقسیم کنید. حدود 0.2 ثانیه سکوت در اول و آخر هر فایل مناسب است.

در آخر فایل ها را به 16 bits و ریت 16 kHz تبدیل کنید.