

Wie Applikationen das Sprechen lernen



© AdobeStock_407758537

Sicherheit, Barrierefreiheit, leichtere Handhabung und Steigerung der Effizienz: Sind Geräte in der Lage mit ihren Nutzern über Sprache zu kommunizieren, hat das viele Vorteile. Meldet ein Warnsystem im Krankenhaus über Lautsprecher zum Beispiel das konkrete Szenario beim Patienten, können bereits auf dem Weg dorthin erste Maßnahmen eingeleitet werden. Auch in weitläufigen Produktionsstätten ist eine Implementierung von Sprache eine sinnvolle Ergänzung, um Abläufe zu optimieren. Wichtig dabei ist in jedem Fall die Verwendung von klarer Sprache, die eindeutig zu verstehen ist.



Autoren:
Stephan Menze (Bild),
Head of Global Innovation
Management und
Marvin Schiffel,
Technischer Support
Rutronik Elektronik
Bauelemente GmbH
www.rutronik.com/de

Text-To-Speech-Software

Für eine qualitativ hochwertige Sprachausgabe wurden bislang individuelle Studioaufnahmen mit professionellen Sprechern vorgenommen, was ein äußerst kosten-

und zeitintensives Vorgehen ist. Einen wirtschaftlich interessanten Lösungsansatz realisiert die Nutzung einer spezialisierten Text-To-Speech-Software für die Phrasengenerierung und eines entsprechenden Chips. Dieser wurde auf ein eigens dafür angepasstes Adapter Board integriert. Dadurch wird es möglich über eine SPI-Schnittstelle mit z. B. Mikrocontrollern zu kommunizieren.

Der folgende Text zeigt eine Step-by-Step-Anleitung für die Erstinutzung des Text-To-Speech-Adapter Boards:

Das Board (Bild 1) eignet sich optimal für Human Machine Interface (HMI)-Anwendungen, die Sprachführung und Audiowiedergabe erfordern, wie sie vor allem im Smart Home, bei Sicherheits- und Alarmanlagen für das indus-



Bild 1: Der Epson Chip ASIC S1V3G340 wurde auf dem Adapter Board Text To Speech von Rutronik integriert

trielle wie private Umfeld, medizinischen oder Fitnessgeräten, sowie Lernanwendungen, wie z. B. Übersetzer, zu finden sind.

Herzstück

des Adapter Boards ist ein neu entwickelter Chip. Er wird vom Host-Mikrocontroller gesteuert

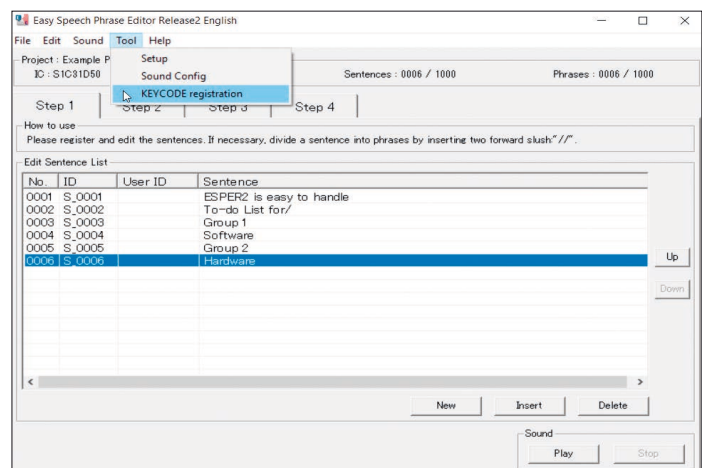


Bild 3: Keycode Registrierung: Auswahl der Schaltfläche „Tool“ und weiter „KEYCODE-registration“

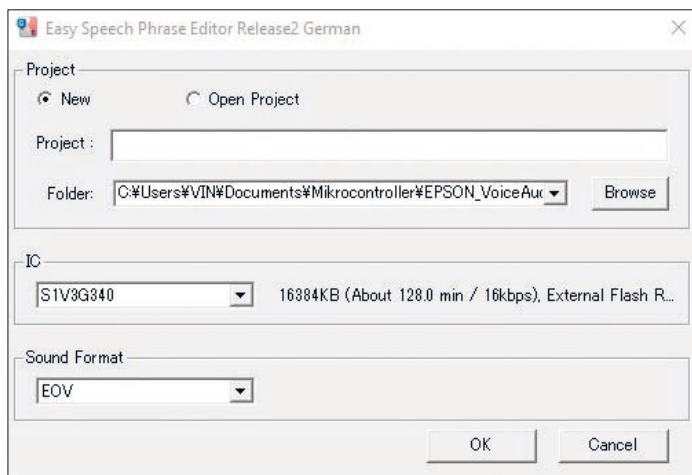


Bild 2: Start der ESPER2 Software und Erstellen eines neuen Projektes mit dem verwendeten IC/MCU. Die für die Programmierung benötigten Software-Pakete sind das Epson ESPER2 (Easy Speech Phrase Editor Release 2) und das RutEpsonFlashTool

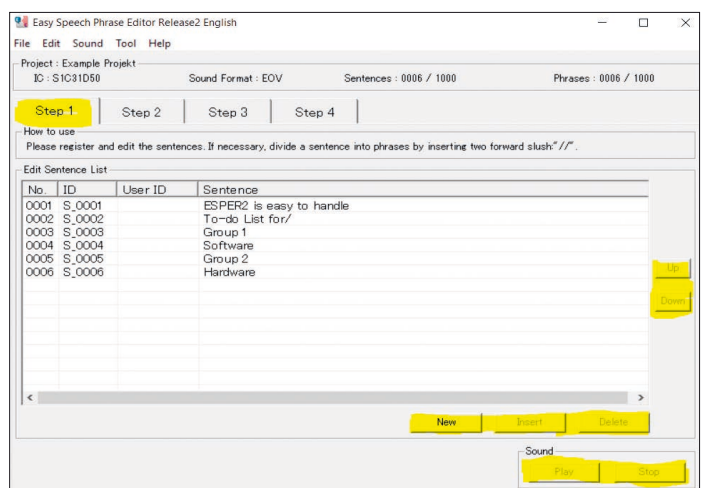


Bild 4: Schritt 1 – Einen neuen Satz anlegen

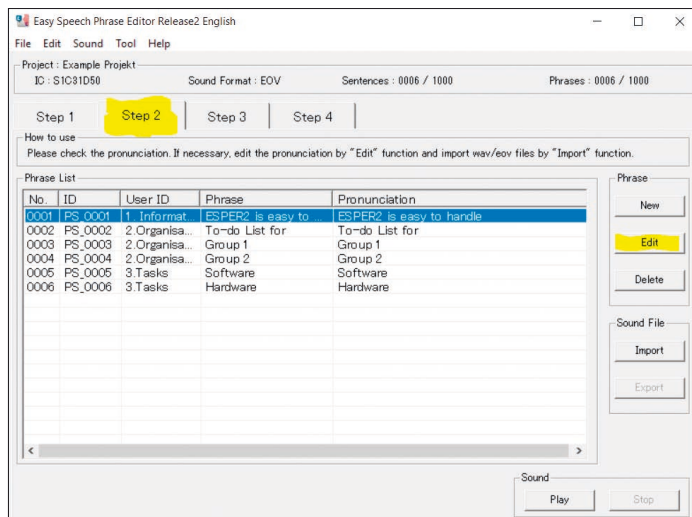


Bild 5: Schritt 2 – Über die „Import“-Schaltfläche können auch Tondateien im wav- oder EOv-Format eingefügt werden

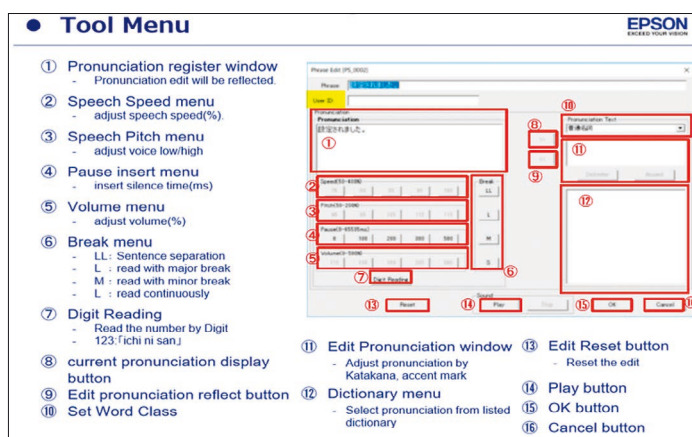


Bild 6: Fenster für Ausspracheoptimierung

und kann zuvor definierte Sprache, die sich als Binary im externen NOR-Flash befindet, abspielen. Die USB-to-SPI-Bridge wandelt beim Flashvorgang die Daten vom USB-Protokoll in ein serielles Protokoll um.

Der Chip unterstützt in Kombination mit dem Voice Data Creation PC Tool zwölf Sprachen: Amerikanisches/britisches Englisch, Chinesisch, Japanisch, Koreanisch, (amerikanisches) Spanisch, (kanadisches) Französisch, Deutsch, Italienisch und Russisch. Vorformulierte Sätze können als Text im CSV-Format in das Tool importiert oder in eine Editormaske direkt eingegeben werden. Mit Hilfe des Tools wird daraus eine Sprachdatei erzeugt. Für die korrekte und natürliche Aussprache und Betonung analysiert das Tool auch den Satzbau der Texte. Zudem verfügt es über ein umfangreiches Wörterbuch. Die Aussprache von Produkt-

und Eigennamen oder Wortschöpfungen, die nicht im Wörterbuch hinterlegt sind, lässt sich mit der Editierfunktion in die gewünschte Form bringen. Dadurch ist die Qualität der generierten Audiodateien vom gesprochenen Wort eines Menschen kaum zu unterscheiden. Zusätzlich sind mittels des kostenlosen Sprachsynthese-IDE-Tools Aufnahmen eigener Sprachführungsaudios in Studioqualität möglich.

1. Vorbereitung auf Software-Ebene

Zum Programmieren und Laden des Adapter Boards werden zwei verschiedene Softwarepakete benötigt. Alle erzeugten Sprach- und Audiodaten werden mit einer Nummer versehen, die in ein Register des Hardware-Prozessors geschrieben wird, der dann die jeweiligen Sprach- und Audiodaten abspielt (Bild 2). Eine zusätzliche Software zur Verknüpfung

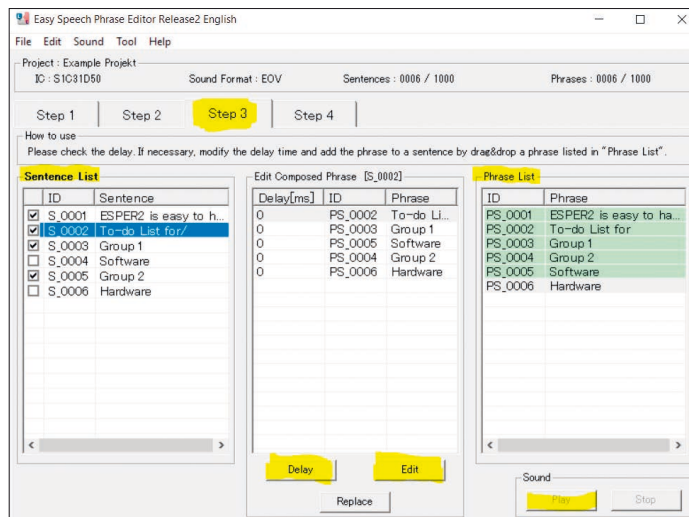


Bild 7: Schritt 3 – Kombinieren von Satzpaaren: Die Sentence List besteht aus allen Sätzen die entweder als CSV-Datei hochgeladen wurden oder selbstständig im Tool generiert wurden

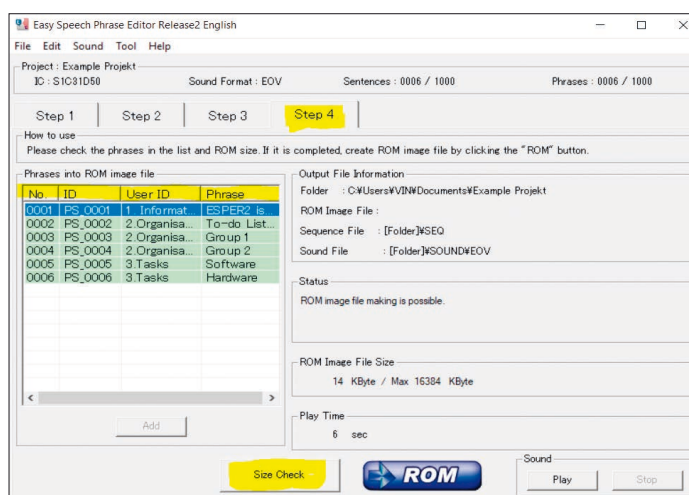


Bild 8: Schritt 4 – Überprüfen auf Vollständigkeit

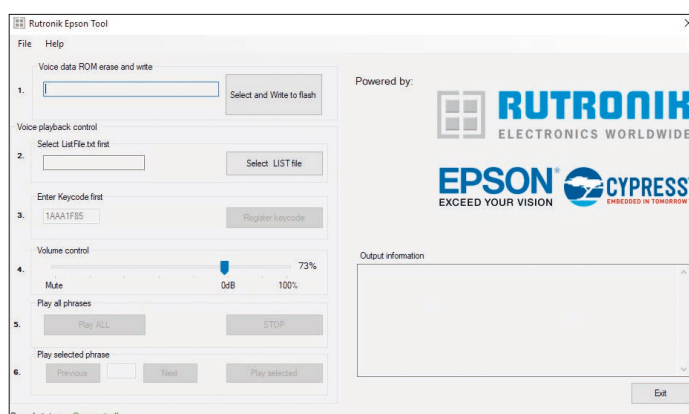


Bild 9: Start-Ansicht des Rutronik Tools

fung der Sprach- und Audiodaten ist damit nicht nötig. Bevor eigene Projekte über die Software realisiert werden können muss die Software über einen Verifizierungscode freigeschaltet werden (Bild 3).

Was soll die Anwendung sagen?

Über die Schaltfläche „New“, oder „Insert“ lässt sich ein neuer Satz anlegen (Bild 4). Die Reihenfolge wird über die Funktionen „Up“ und

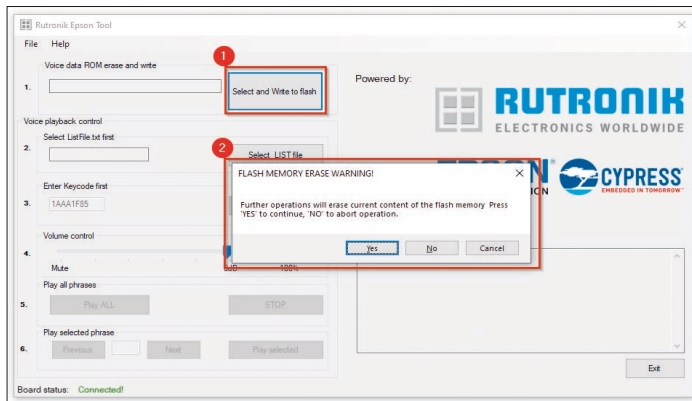


Bild 10: Rutronik Tool Schritt 1 - Bevor neue Daten in das Board geladen werden können, ist es nötig, das Flash zu löschen

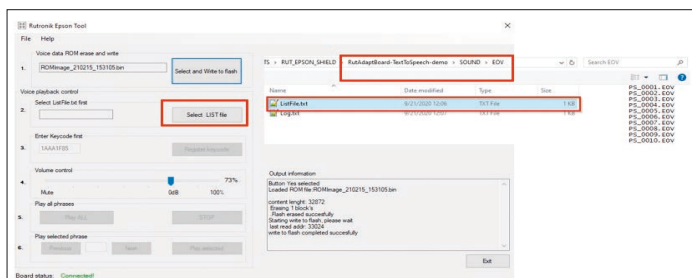


Bild 11: Rutronik Tool Schritt 2

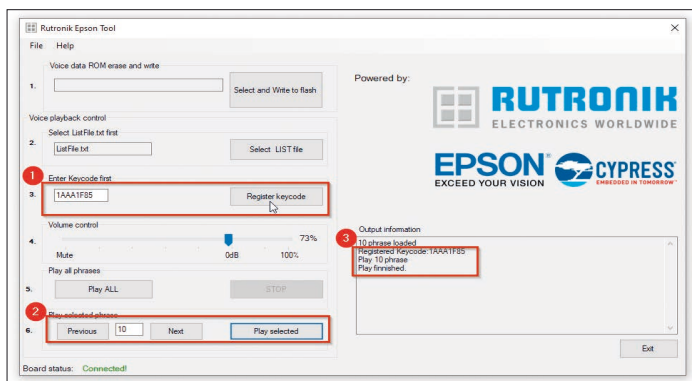


Bild 12: Rutronik Tool Schritt 3 – Nach den drei Schritten kommt der PSoC zum Einsatz

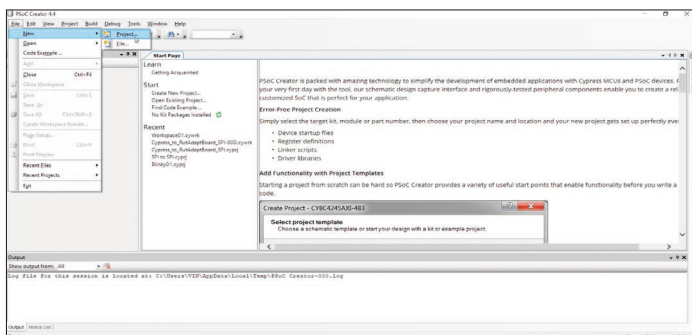


Bild 13: Projektumsetzung für den Cypress PSoC-Controller im PSoC Creator: „File“ -> „New“ -> „Project“

„Down“ verändert. Zur Ergebniskontrolle kann der erstellte gesprochene Text über den PC-Lautsprecher abgespielt werden („Play“ bzw. Stop“).

Es ist anschließend möglich, die Aussprache einzelner Sätze und Wörter zu bearbeiten (Bild 5). Dazu wird der ausgewählte Bereich ausgewählt und durch Klick auf „Edit“ in

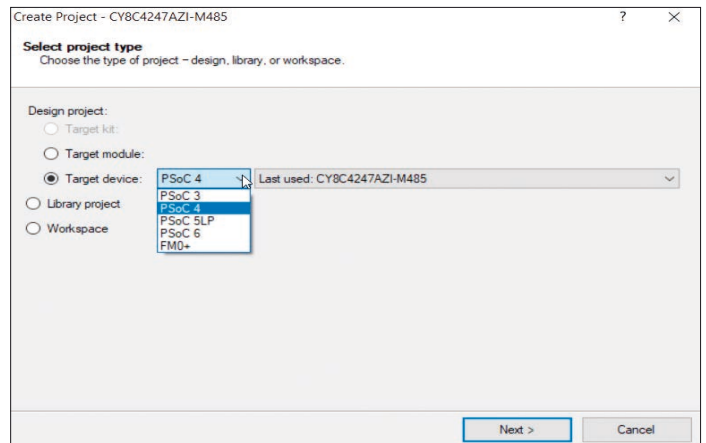


Bild 14: Auswahl des Ziel-Controllers

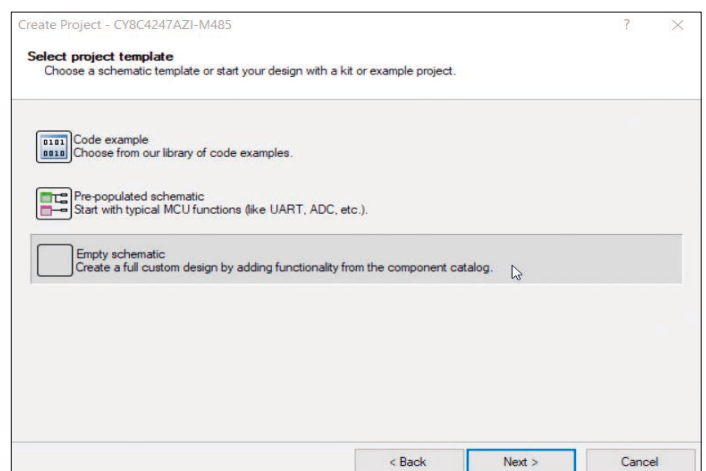


Bild 15: Alternativ ist es auch möglich, mit vorgefertigten Code-Beispielen weiterzufahren

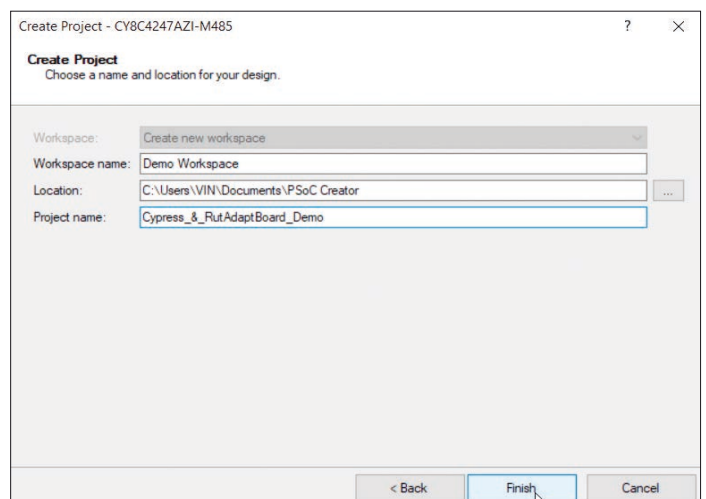


Bild 16: Bildschirmansicht zur Benennung des Workspaces und des Projektes

einem Bearbeitungsfenster (Bild 6) geöffnet.

Nun lassen sich diverse Punkte verändern: So können Aufzeichnungen durch die Verwendung der „User ID“ kategorisiert oder Abspielgeschwindigkeit, Tonhöhe, Pausen, Lautstärke usw. optimiert werden.

Die Veränderungen können zur Kontrolle abgespielt und ggf. weiter angepasst werden. Mit einem finalen „OK“ erfolgt die endgültige Bestätigung.

Mittels einfacher „Drag & Drop“-Funktion werden nun Satzpaare gebildet. Der zu bearbeitende

	RutAdaptBoard	Function	PSoC Port	Plug-in board
P 0.0	P 6.4	MessageReady	P 1.0	LED1
P 0.1	P 6.5	StandByExit	P 1.1	LED2
P 0.2	P 6.7	DeviceReset	P 1.2	LED3
P 0.3	P 1.2	Mute	P 1.3	LED4
P 0.4	P 1.4	SPI:Mosi	P 1.4	Button
P 0.5	P 1.5	SPI:Miso		
P 0.6	P 1.6	SPI:SCLK		
P 0.7	P 6.8	SPI:SS		

Tabelle 1: Übersicht der Anschlüsse

Satz wird dafür aus der „Sentence List“ (links in Bild 7) ausgewählt, die gewünschten Phrasen aus der „Phrase List“ (rechts) hinzugefügt. Die Phrasendateien können ebenfalls individuell angepasst werden. Außerdem ist es möglich Audio-dateien, wie z. B. eine spezielle Wartemusik, zu dem Satzkonstrukt zu ergänzen. Für einen natürlicheren Sprachfluss sind auch Pausen zwischen den Fragmenten wichtig, diese werden mittels der Funktion „Delay“ integriert.

Vollständigkeit prüfen

Im nächsten Schritt werden die geplanten Sätze auf Vollständigkeit überprüft (Bild 8). Die einzelnen Satzteile können dafür nach „Nr.“, „ID“, „User ID“, „Phrase“ sortiert werden. Nach der Überprüfung empfiehlt es sich, die Größe der zu konvertierenden Datei mit „Size Check“ zu kontrollieren, um unnötige Anforderungen an die Speicherkapazitäten zu vermeiden, die

die Funktionsweise der Anwendung beeinträchtigen könnten. Durch den abschließenden Klick auf „ROM“ wird die Sounddatei final erstellt.

So bekommt die Anwendung ihre Stimme

Nach der Definition, welche Texte über die geplante Anwendung ausgegeben werden sollen, wird die erstellte Sounddatei auf das Adapter-Board geladen. Dazu ist eine Verbindung des Boards über ein USB-Kabel mit dem PC, sowie über ein Aux-Kabel mit den Lautsprechern nötig. Wichtig dabei ist, dass die Schalterstellungen auf „ON-USB“ stehen.

Nun können die erstellten Sounddateien geladen werden (Bild 9). Dafür muss zunächst über den Bereich „Select and Write to Flash“ -> „SOUND“ -> „EOV“ das gewünschte Projekt ausgewählt werden (Bild 10).

Im Anschluss wird unter „Select LIST file“ (Bild 11) die Datei ausgewählt, die die Software anzeigen und durch die Sätze navigieren soll.

Änderungen abschließend verifizieren

Um alle Änderungen abschließend zu verifizieren, erfolgt die Registrierung durch einen Keycode. Dabei handelt es sich um denselben Code, der zum Freischalten verwendet wurde. Erst dann ist es möglich die importierten Sounddateien mit Hilfe der Tasten „Play selected“, „Next“ und „Previous“ anzuhören (Bild 12).

Auf Kommando sprechen lernen

Nach den Einstellungen im Adapter Board, folgt die Projektumsetzung für den Controller. Als erstes wird dieser über USB mit dem PC verbunden. Im Mikrocontroller Creator erfolgt der erste softwareseitige Schritt, das Anlegen eines eigenen Arbeitsbereichs (Bild 13) und anschließend die Auswahl des Ziel-Controllers (Bild 14):

Für ein komplett individualisiertes Design wird im nächsten Fenster die Auswahl „Empty Schematic“ getroffen (Bild 15). Anschließend erfolgt die Benennung des Projektes und des Arbeitsbereiches (Bild 16).

Alle wichtigen Treiberdateien, die für die Kommunikation zwischen Controller und Anwendung nötig sind, sind in einer ZIP-Datei zusammengestellt und werden als exklusive Serviceleistung für Kunden auf Nachfrage angeboten. Um die opti-

male Funktionalität zu gewährleisten, dürfen hier keine Änderungen vorgenommen werden (Bild 17 und 18).

Nach der Software kommt die Hardware-Konfiguration und Handhabung

Neben dem Adapter Board und dem Controller sind zudem ein Lautsprecherkabel, sowie der passende Lautsprecher, ein USB-Micro-B-Kabel mit Datenübertragungsfunktion und eine Stromquelle (z. B. der Laptop) erforderlich. Das verwendete Beispiel nutzt ein spezielles Prototyping Kit mit vier externen LEDs als Statusregister und einer externen Taste als Eingangssignal. Zunächst werden die beiden Boards miteinander und mit der Peripherie verbunden (Bild 19).

Nachdem der Controller über USB mit einer Stromquelle verbunden wurde, kann die Audiowiedergabe per Knopfdruck gestartet werden.

Damit Reden Gold wird

Die Time-To-Market ist bei der Entwicklung von Anwendungen der letztlich entscheidende Punkt: Langwierige Vorarbeiten und zusätzliche Ausgaben wie z. B. Studioaufnahmen für die Sprachausgabe eines Smart Devices könnten ein Projekt unrentabel werden lassen, so dass es gar nicht erst in die Serienproduktion geht. Wer die Entwicklungszeit für eine hochwertige Sprachausgabe nochmal deutlich verkürzen möchte, greift deshalb am besten auf ein Arduino-kompatibles Adapter-Board (Arduino Shield) zurück. Es kann auf jedes handelsübliche Mikrocontroller-Evaluation-Kit mit Arduino-Interface gesteckt werden. ◀

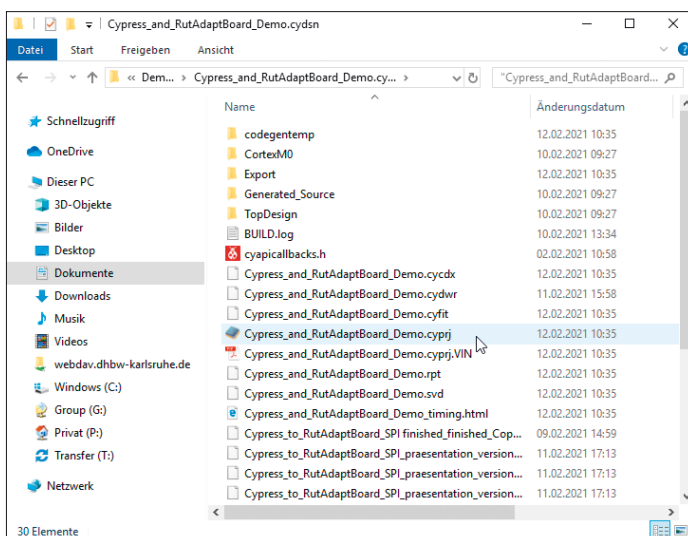


Bild 17: Nach dem Entpacken des Ordners in den Workspace wird die Applikation „Cypress_and_RutAdaptBoard_Demo.cypri“ gestartet

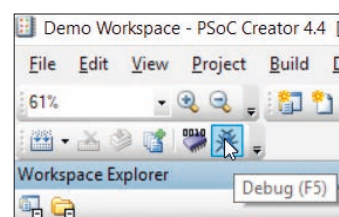


Bild 18: Schaltflächen zu Debuggen und Flashen befinden sich in der oberen linken Ecke

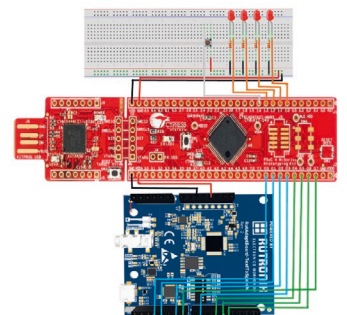


Bild 19: So werden das Rutronik Adapter Board – Text To Speech, der Cypress Controller und das PSoC Prototyping Kit miteinander verbunden